

QY	2218	TAATTTTATGCGTGAACACAGAGTTTGGCGACATTTGGACCATCAGAAAGCAGAAGTGTGCAC	2277
Db	783	TAATTTTATGCGTGAACACAGAGTTTGGCGACATTTGGACCATCAGAAAGCAGAAGTGTGCAC	842
QY	2278	TATTTCAAGTCAGTGTCTCAAAAGTTTTCAGATGTTAAAGCTGGTGTGATGCAGTTTCATGCCAG	2337
Db	843	TATTTCAAGTCAGTGTCTCAAAAGTTTTCAGATGTTAAAGCTGGTGTGATGCAGTTTCATGCCAG	902
QY	2338	TGATCCGAGTACTTTTGGGAAGCCACAGACAGGTGGATCTCTTGAGCCCGAGGAGTTTGAGGC	2397
Db	903	TGATCCGAGTACTTTTGGGAAGCCACAGACAGGTGGATCTCTTGAGCCCGAGGAGTTTGAGGC	962
QY	2398	CAGACTGCACAAACACAGTGTGAGACTCGTTTCTTACAAATAATTAATAAATTTAGCCAGGTGT	2457
Db	963	CAGACTGCACAAACACAGTGTGAGACTCGTTTCTTCAAAATAATTAATAAATTTAGCCAGGTGT	1022
QY	2458	GGTGGTGCACACCTGTAGTCCGAGTACTCAGGAGGCTGAGG-TAGTAGGATGTTTGGAG	2516
Db	1023	GGTGGTGCACACCTGTAGTCCGAGTACTCAGGAGGCTGAGG-TAGTAGGATGTTTGGAG	1082
QY	2517	ACTGGGAGGTTGAGGCTGAACTCAGCCAGGATCTTGCCACCACA	2560
Db	1083	ACTGGGAGGTTGAGGCTGAACTCAGCCAGGATCTTGCCACCACA	1126
RESULT 15			
CQ494252		1128 bp	DNA
LOCUS		Sequence 26119 from Patent WO0160860.	linear
DEFINITION			PAT 30-JAN-2004
ACCESSION		CQ494252	
VERSION		CQ494252.1	GI:41459871
KEYWORDS			
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1	Schlegel R., Endege, W.O. and Monahan, J.E.
AUTHORS			Genes differentially expressed in human prostate cancer and their
TITLE		use	
JOURNAL		Patent: WO 0160860-A 26119 23-AUG-2001;	
		Millennium Predictive Medicine, Inc. (US)	
FEATURES		Location/Qualifiers	
source		1..1128	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Query Match		28.4%;	Score 1068; DB 6; Length 1128;
Best Local Similarity		99.0%;	Pred. No. 1.2e-176;
Matches 1102;		Conservative 0;	Mismatches 20; Indels 2; Gaps 2;
QY	1438	CCAAATTTCAAATGCCAACATCTCCCATTCCTTACCTATTTTCATCTTTGAGGTGTAAT	1497
Db	4	CTACAAATTTAAATGCAACATCTGCCATCCAA-TACCTATTTTCATCTTTGAGGGTAAT	62
QY	1498	CTACTCAATAAAGTGTGTAGACACAGTGCACAGACCCCTTTGCTTAACCTGACATTTACTTC	1557
Db	63	CTACTCAATAAAGTGTGTAGACACAGTGCACAGACCCCTTTGCTTAACCTGACATTTACTTC	1222
QY	1558	AATTTTCTTTTCTATGTAAGTGTGATATTTTTCATATAAACTTGCAGTAATAGTTTCAAA	1617
Db	123	AATTTTCTTTTCTATGTAAGTGTGATATTTTTCATATAAACTTGCAGTAATAGTTTCAAA	162
QY	1618	AATTAATAGTTTTCATGATTTGGCTTTTCTCGAAGAGAGAAAATGGAAGTGTCACAAAATAA	1677
Db	183	AATTAATAGTTTTCATGATTTGGCTTTTCTCGAAGAGAGAAAATGGAAGTGTCACAAAATAA	242
QY	1678	AAAAGATGAATGAAGCATATATAATTTGTCATATTTTTCATTTTCTAGTCAACAGAGA	1737
Db	243	AAAAGATGAATGAAGCATATATAATTTGTCATATTTTTCATTTTCTAGTCAACAGAGA	302

Search completed: November 18, 2004, 18:25:07
Job time : 15382 secs

[illegible]

```
Db      1083 ACTGGAGGTTGAGGCTGAAGTGAAGCCAGGATCTTGGCCACCACA 1126

RESULT 12
CQ488563
LOCUS      CQ488563                1128 bp      DNA      linear      PAT 30-JAN-2004
DEFINITION Sequence 20430 from Patent WO0160860.
ACCESSION  CQ488563
VERSION    CQ488563.1 GI:41454182
KEYWORDS   .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS    Schlegel R., Endege W.O. and Monahan J.E.
TITLE      Genes differentially expressed in human prostate cancer and their
            use
JOURNAL    Patent: WO 0160860-A 20430 23-AUG-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES   source
            Location/Qualifiers
            1..1128
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match      28.4%; Score 1068; DB 6; Length 1128;
Best Local Similarity 98.0%; Pred. No. 1.2e-176;
Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

Qy      1438 CCAATTTCAATGCGACATCTCCCATCCAATTACCTATTTCATCTTTGAGGTGAAT 1497
Db      4 CTACAAATTTAAATGCAACATCTGCCATCAA-TACCTATTTCATCTTTGAGGGTAAT 62

Qy      1498 CTACTCAATAAATGTAAGCAGTGCACAGACCCCTTTCCTAACCTGCATTTACTTTC 1557
Db      63 CTACTCAATAAATGTAAGCAGTGCACAGACCCCTTTCCTAACCTGCATTTACTTTC 122

Qy      1558 AATTTTCTTTTCTATGCTAGTATTTTTCATATATAAATCTTGCAATAGTGTCAAA 1617
Db      123 AATTTTCTTTTCTATGCTAGTATTTTTCATATATAAATCTTGCAATAGTGTCAAA 182

Qy      1618 AATTAATAGTTTTGACATGGCTTTCTGAGAGAGAAATGAAAGTGCACAAAATAA 1677
Db      183 AATTAATAGTTTTGACATGGCTTTCTGAGAGAGAAATGAAAGTGCACAAAATAA 242

Qy      1678 AAAAGATGAATGAAGCATATATAATTTGTCAAATTTTCAATTTTCTAGTCAACAGAGA 1737
Db      243 AAAAGATGAATGAAGCATATATAATTTGTCAAATTTTCTAGTCAACAGAGA 302

Qy      1738 ATCGAAGATTCTGTTCAAAATATTAGTAAATGAAATGAAATGAAATGAAATGAAATGAA 1797
Db      303 ATCGAAGATTCTGTTCAAAATATTAGTAAATGAAATGAAATGAAATGAAATGAAATGAA 362

Qy      1798 TTTCGACACACTAGTTAAATTTAACTGTGACTAGTATCTCTACCGAAGTGAATGTT 1857
Db      363 TTTCGACACACTAGTTAAATTTAACTGTGACTAGTATCTCTACCGAAGTGAATGTT 422

Qy      1858 AGTTTCTGGTTTAAATTTCAAGCAAACTGAAATAATCCATCTTAATTTATGCTTTCTTT 1917
Db      423 AGTTTCTGGTTTAAATTTCAAGCAAACTGAAATAATCCATCTTAATTTATGCTTTCTTT 482

Qy      1918 CCCAAGAAGTTTTTAAATGATATGATGCGAGCTTCTTAATTTGGAGACAAAGCCTTAATGA 1977
Db      483 CCCAAGAAGTTTTTAAATGATATGATGCGAGCTTCTTAATTTGGAGACAAAGCCTTAATGA 542

Qy      1978 CAATGCATTCATATATATATTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG 2037
Db      543 CAATGCATTCATATATATATTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG 602

Qy      2038 ATGAGATGCTTGGGACCAAGAGTGTTTTGGATTTTTCAGATTTTATTTTGGATTTTGGATA 2097
Db      602 ATGAGATGCTTGGGACCAAGAGTGTTTTGGATTTTTCAGATTTTATTTTGGATTTTGGATA 1617

Db      603 ATGAGATGCTTGGACCCAGAAAGTCTTTTGGATTTTCAGATTTTATTTTGGATTTTGGATA 662
Qy      2098 TTTCCATACATATAAATGAGAGAGTTGAAAATGGGATTCAGTCTAATATCAATAAATTCAC 2157
Db      663 TTTCCATACATATAAATGAGAGAGTTGAAAATGGGATTCAGTCTAATATCAATAGATGCAC 722
Qy      2158 TTATGTTTGATATACACCTTATCTGAATAGCCTGAAGGTAAATTTTATACAAATATTTTAAA 2217
Db      723 TTATGTTTGATATACACCTTATCTGAATAGCCTGAAGGTAAATTTTATACAAATATTTTAAA 782
Qy      2218 TAATTTTATGCTGAAACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAAGTGTAC 2277
Db      783 TAATTTTATGCTGAAACAGAGTTTTCGCAATTTGGACCATCAGAAAGCAAGTGTAC 842
Qy      2278 TATTTCAAGTCAGTGTCTCAAAAAGTTTCAGATGTTAAAGCTGGTGATGCAGTTTCATGCCAG 2337
Db      843 TATTTCAAGTCAGTGTCTCAAAAAGTTTCAGATGTTAAAGCTGGTGATGCAGTTTCATGCCAG 902
Qy      2338 TGATCCAGTACITTTGGAGAGCCAGACAGGTGATCTCTTGAGCCAGAGTTTGAGGC 2397
Db      903 TGATCCAGTACITTTGGAGAGCCAGACAGGTGATCTCTTGAGCCAGAGTTTGAGGC 962
Qy      2398 CAGACTCCACAACACAGTGAACCTCTCTTCTCAAAATAAATTTAAATAATTTAGCCAGTGT 2457
Db      963 CAGACTCCACAACACAGTGAACCTCTCTTCTCAAAATAAATTTAAATAATTTAGCCAGTGT 1022
Qy      2458 GGTGGTGACACCTGTAGTCTCCAGGTACTCAGGAGGCTGAGG-TAGTAGGATTTTGTAG 2516
Db      1023 GGTGGTGACACCTGTAGTCTCCAGGTACTCAGGAGGCTGAGGTTATTTGGGATTTTGTAG 1082
Qy      2517 ACTGGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGGCCACCACA 2560
Db      1083 ACTGGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGGCCACCACA 1126

RESULT 13
CQ488741
LOCUS      CQ488741                1128 bp      DNA      linear      PAT 30-JAN-2004
DEFINITION Sequence 20608 from Patent WO0160860.
ACCESSION  CQ488741
VERSION    CQ488741.1 GI:41454360
KEYWORDS   .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS    Schlegel R., Endege W.O. and Monahan J.E.
TITLE      Genes differentially expressed in human prostate cancer and their
            use
JOURNAL    Patent: WO 0160860-A 20608 23-AUG-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES   source
            Location/Qualifiers
            1..1128
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match      28.4%; Score 1068; DB 6; Length 1128;
Best Local Similarity 98.0%; Pred. No. 1.2e-176;
Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

Qy      1438 CCAATTTCAATGCGACATCTCCCATCCAATTACCTATTTCATCTTTGAGGTGAAT 1497
Db      4 CTACAAATTTAAATGCAACATCTGCCATCAA-TACCTATTTCATCTTTGAGGGTAAT 62

Qy      1498 CTACTCAATAAATGTAAGCAGTGCACAGTGCACAGACCCCTTTCCTAACCTGCATTTACTTTC 1557
Db      63 CTACTCAATAAATGTAAGCAGTGCACAGTGCACAGACCCCTTTCCTAACCTGCATTTACTTTC 122

Qy      1558 AATTTTCTTTTCTATGCTAGTATTTTTCATATATAAATCTTGCAATAGTGTCAAA 1617
Db      123 AATTTTCTTTTCTATGCTAGTATTTTTCATATATAAATCTTGCAATAGTGTCAAA 182

Qy      1618 AATTAATAGTTTTGACATGGCTTTCTGAGAGAGAAATGAAAGTGCACAAAATAA 1677
Db      183 AATTAATAGTTTTGACATGGCTTTCTGAGAGAGAAATGAAAGTGCACAAAATAA 242

Qy      1678 AAAAGATGAATGAAGCATATATAATTTGTCAAATTTTCAATTTTCTAGTCAACAGAGA 1737
Db      243 AAAAGATGAATGAAGCATATATAATTTGTCAAATTTTCTAGTCAACAGAGA 302

Qy      1738 ATCGAAGATTCTGTTCAAAATATTAGTAAATGAAATGAAATGAAATGAAATGAAATGAA 1797
Db      303 ATCGAAGATTCTGTTCAAAATATTAGTAAATGAAATGAAATGAAATGAAATGAAATGAA 362

Qy      1798 TTTCGACACACTAGTTAAATTTAACTGTGACTAGTATCTCTACCGAAGTGAATGTT 1857
Db      363 TTTCGACACACTAGTTAAATTTAACTGTGACTAGTATCTCTACCGAAGTGAATGTT 422

Qy      1858 AGTTTCTGGTTTAAATTTCAAGCAAACTGAAATAATCCATCTTAATTTATGCTTTCTTT 1917
Db      423 AGTTTCTGGTTTAAATTTCAAGCAAACTGAAATAATCCATCTTAATTTATGCTTTCTTT 482

Qy      1918 CCCAAGAAGTTTTTAAATGATATGATGCGAGCTTCTTAATTTGGAGACAAAGCCTTAATGA 1977
Db      483 CCCAAGAAGTTTTTAAATGATATGATGCGAGCTTCTTAATTTGGAGACAAAGCCTTAATGA 542

Qy      1978 CAATGCATTCATATATATATTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG 2037
Db      543 CAATGCATTCATATATATATTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG 602

Qy      2038 ATGAGATGCTTGGGACCAAGAGTGTTTTGGATTTTTCAGATTTTATTTTGGATTTTGGATA 2097
Db      602 ATGAGATGCTTGGGACCAAGAGTGTTTTGGATTTTTCAGATTTTATTTTGGATTTTGGATA 1617
```

Db	1301	
----	------	--

Db	4	CTACAAATTTAAATGCAACATCTGCCCATCCAA-TACCTATTTTCATCTTTGAGGGTAAT	62
Qy	1498	CTACTCAATAAATGTTGTAAGACCGAGTGAACCGCTTTGCTAACTGACATTTACTTC	1557
Db	63	CTACTCAATAAATGTTGTAAGACCGAGTGAACCGCTTTGCTAACTGACATTTACTTC	132
Qy	1558	AAATTTTCTTTTCTATGTAAGTATTTTGCATATAAACTTGACAGTAATAGTTCAAA	1617
Db	123	AAATTTTCTTTTCTATGTAAGTATTTTGCATATAAACTTGACAGTAATAGTTCAAA	182
Qy	1618	AAATTAATAGTTTGTGACATTTGGCTTTCTGAGAAGAGAAATTCGAAAGTGTCAAAAATA	1677
Db	183	AAATTAATAGTTTGTGACATTTGGCTTTCTGAGAAGAGAAATTCGAAAGTGTCAAAAATA	242
Qy	1678	AAAAAGATGAATGAAGCATATATATATTTTCAATTTTCTAGTCAACAGAGA	1737
Db	243	AAAAAGATGAATGAAGCATATATATATTTTCAATTTTCTAGTCAACAGAGA	302
Qy	1738	ATCGAAGGATCTGTTTCAAAATATTAGTAAATAATTTGAAATTAATCTTGTGCTTATTTTG	1797
Db	303	ATCGAAGGATCTGTTTCAAAATATTAGTAAATAATTTGAAATTAATCTTGTGCTTATTTTG	362
Qy	1798	TTTGCAACACACTAGTTAAATTTAACTGTGACTAGTTATCTCTACCGAAGGTGAGTGT	1857
Db	363	TTTGCAACACACTAGTTAAATTTAACTGTGACTAGTTATCTCTACCGAAGGTGAGTGT	422
Qy	1858	AGTTTCTGTTTAAATTTCAAGCAAACTGGAATAATTAATCCATCTAATTTGCTTTCTTT	1917
Db	423	AGTTTCTGTTTAAATTTCAAGCAAACTGGAATAATTAATCCATCTAATTTGCTTTCTTT	482
Qy	1918	CCCAAGAGTTTAAATTTAACTGATGCGCAGCTTCTCTAAATTTGGAGACAAAGCCCTTAATTGA	1977
Db	483	CCCAAGAGTTTAAATTTAACTGATGCGCAGCTTCTCTAAATTTGGAGACAAAGCCCTTAATTGA	542
Qy	1978	CAATGCAATTCATTTATATATTTTGTATAGTTTACAGTATACGAGTTGAGTATCCCTTAG	2037
Db	543	CAATGCAATTCATTTATATATTTTGTATAGTTTACAGTATACGAGTTGAGTATCCCTTAG	602
Qy	2038	ATGAGATGCTTGGGACCAAGTGTGTTTGGATTTTCAAGTTTATTTTGGATTTTGGATTA	2097
Db	603	ATGAGATGCTTGGGACCAAGTGTGTTTGGATTTTCAAGTTTATTTTGGATTTTGGATTA	662
Qy	2098	TTTCCATACATATAATGAGAGGTTTGGAAAATCGGATTCAGTCTAATCATATAAATTCAC	2157
Db	663	TTTCCATACATATAATGAGAGGTTTGGAAAATCGGATTCAGTCTAATCATATAAATTCAC	722
Qy	2158	TTATGTTTGTATATACACCTTTATCTGAATAGCCTGAAGGTAATTTTATACAAATTTTAA	2217
Db	723	TTATGTTTGTATATACACCTTTATCTGAATAGCCTGAAGGTAATTTTATACAAATTTTAA	782
Qy	2218	TAAATTTATGCTGAAAACAGAGTTTGGGACCATTTGGACCATCAGAAAGAGAGTGTAC	2277
Db	783	TAAATTTATGCTGAAAACAGAGTTTGGGACCATTTGGACCATCAGAAAGAGAGTGTAC	842
Qy	2278	TATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATCTTAAAGCTGTTGATGTCAGTTCATGCCAG	2337
Db	843	TATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATCTTAAAGCTGTTGATGTCAGTTCATGCCAG	902
Qy	2338	TGATCCGAGTACTTTGGGAAAGCAACAGAGGTGATCTCTTTGAGCCCAAGGTTTGAAGC	2397
Db	903	TGATCCGAGTACTTTGGGAAAGCAACAGAGGTGATCTCTTTGAGCCCAAGGTTTGAAGC	962
Qy	2398	CAGACTGCAACACAGAGTGTGCTGTTTCTCAATTAATTAATAAATTTAGCCAGGTGT	2457
Db	963	CAGACTGCAACACAGAGTGTGCTGTTTCTCAATTAATTAATAAATTTAGCCAGGTGT	1022
Qy	2458	GGTGTGTCACACCTGTAGTCCCAAGGTACTCAGGAGGTGAGG-TAGTAGGATTTTGTAG	2516
Db	1023	GGTGTGTCACACCTGTAGTCCCAAGGTACTCAGGAGGTGAGG-TAGTAGGATTTTGTAG	1082
Qy	2517	ACTGGAGGTTGAGGCTGAATCAGGACAGGATCTTGGCCACCA 2560	

RESULT 10	AB060908	1924 bp	mRNA	linear	PRI 19-MAY-2004
LOCUS	Macaca fascicularis brain cDNA clone:Qtra-14007, full insert				
DEFINITION	sequence.				
ACCESSION	AB060908				
VERSION	AB060908.1	GI:13874585			
KEYWORDS	oligo capping; f1s (full insert sequence).				
SOURCE	Macaca fascicularis (crab-eating macaque)				
ORGANISM	Macaca fascicularis				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;				
	Cercopithecinae; Macaca.				
REFERENCE	1				
AUTHORS	Osada, N., Hida, M., Kusuda, J., Tanuma, R., Hirata, M., Hirai, M.,				
	Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.				
TITLE	Prediction of unidentified human genes on the basis of sequence				
	similarity to novel cDNAs from cynomolgus monkey brain				
JOURNAL	(er) Genome Biol. 3, research0006.1-0006.5 (2001)				
REFERENCE	2	(bases 1 to 1924)			
AUTHORS	Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-APR-2001) Katsuyuki Hashimoto, National Institute of				
	Infectious Diseases, Division of Genetic Resources; 23-1, Toyama				
	1-chome, Shinkuku-ku, Tokyo 162-8640, Japan				
	(E-mail, khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank//,				
	Tel:81-3-5285-1111 (ex.2120), Fax:81-3-5285-1181)				
COMMENT	Lab host: TOP10				
	Vector: pME18S-FL3 (Acc.No. AB009864)				
	R. Site1: DraIII (CACTGTGTG)				
	R. Site2: DraIII (CACCATGTG)				
	Description: 1st strand cDNA was primed with an oligo(dT) primer				
	[ATGTGGCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized				
	using specific 5' and 3' primers and amplified by PCR. The PCR				
	product was digested with SfiI and size selection was performed to				
	exclude fragments <1.5kb. The SfiI-digested PCR product was cloned				
	into distinct DraIII sites of pME18S-FL3. XhoI sites just outside				
	the DraIII sites can be used to isolate the cDNA insert. Libraries				
	were constructed by oligo-capping method				
	(Sugano et al., Institute of Medical Science, University of				
	Tokyo).				
	Custom primer used for sequencing				
	(5' end primer [CTTCTGCTTAAAGCTGCG];				
	3' end primer [CGACCTGCAGCTCGACACA]).				
FEATURES	Location/Qualifiers				
source	1..1924				
	/organism="Macaca fascicularis"				
	/mol_type="mRNA"				
	/db_xref="taxon:9541"				
	/clone="Qtra-14007"				
	/sex="male"				
	/tissue_type="temporal lobe right"				
	/clone_lib="macaque brain cDNA library Qtra"				
	/dev_stage="adult"				
	248..559				
CDS	/codon_start=1				
	/product="hypothetical protein"				
	/protein_id="BAB46903.1"				
	/db_xref="GI:13874586"				
	/translation="MTTERKLQDDKSKVEVSSISNOENENGSGSEVCYTVINHP				
	HRRLSSNDGDYENIDSLTRKVRERERSEYALLRTSVSRPYSCTHEHDVEWLP				
	H"				
ORIGIN					
Query Match	40.0%;	Score 1504.6;	DB 9;	Length 1924;	
Best Local Similarity	92.5%;	Pred. No. 7.3e-253;			
Matches 1616;	Conservative	0;	Mismatches 123;	Indels	7; Gaps 3;
QY	37	CTCCTCGAAAACTCAGTTGCTCGGAGAGAAATCAAAAGAGCCCAAGAAAGGAAACCCA	96		
Db	164	CGCGGAGGACATGGAGTTGCTTCGGAGAGATCAAAAGAGCCCAAGAAAGGAAACCCA	223		

Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1709)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1..1709
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NOVAR2000710"
/tissue_type="ovary"
/clone_lib="NOVAR2"
/dev_stage="adult"
/note="cloning vector: pME18SFL3"

FEATURES
source
44.5%; Score 1673.4; DB 9; Length 1709;
Best Local Similarity 99.9%; Pred. No. 2.6e-282;
Matches 1674; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN
Query Match
1 GAGAACCCGAGTCACCTGTGAAAGATGGGAAATATCTCCTGGGAAACTCAGTTGCCTG 60
35 GAGAACCCGAGTCACCTGTGAAAGATGGGAAATATCTCCTGGGAAACTCAGTTGCCTG 94
61 GAGAGAAATCAAAGAACCCAGAAAGGAAACCCAGATGAGGAAAGAAACCGGAGGAA 120
95 GAGAGAAATCAAAGAACCCAGAAAGGAAACCCAGATGAGGAAAGAAACCGGAGGAA 154
121 ATGACTACATTTGAAAGAAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAG 180
155 ATGACTACATTTGAAAGAAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAG 214
181 ACTTCTAATCAGGAAACGAGAAATGGCAGTGGTCTTGAAGAAAGTGTGCTACACTGTCA 240
215 ACTTCTAATCAGGAAACGAGAAATGGCAGTGGTCTTGAAGAAAGTGTGCTACACTGTCA 274
241 AATCAGATCCCGCATCAGAGATCCTCCTGAGCTCCAAATGATGATGGCTATGAGAACAT 300
275 AATCAGATCCCGCATCAGAGATCCTCCTGAGCTCCAAATGATGATGGCTATGAGAACAT 334
301 GACTCCCTCACAAGGAAAGTGACACAGTTTAGAGAAAGTCAAGACAGAAATATGCCCT 360
335 GACTCCCTCACAAGGAAAGTGACACAGTTTAGAGAAAGTCAAGACAGAAATATGCCCT 394
361 CTTAGGACTCTGTAGTAGGCTGTCTCTGACCCATCAGCATGATTAAGAGTTGTG 420
395 CTTAGGACTCTGTAGTAGGCTGTCTCTGACCCATCAGCATGATTAAGAGTTGTG 454
421 TTTCCACACTAAATCCTCAAGCTGTTTATACCTTCCAGCAATGAAGCAATGAGAA 480
455 TTTCCACACTAAATCCTCAAGCTGTTTATACCTTCCAGCAATGAAGCAATGAGAA 514
481 TAGCAGACTCTGGGAAAGTTTCACTCAGCAGTGCATGAAACATTTCTTCTGGCTA 540
515 TAGCAGACTCTGGGAAAGTTTCACTCAGCAGTGCATGAAACATTTCTTCTGGCTA 574
541 AAGTTTAGAAATATATCTTATATATATATATATATATATATATATATATATATATAT 600
575 AAGTTTAGAAATATATCTTATATATATATATATATATATATATATATATATATATAT 634

QY 601 TGGCTTAGCTGAAATCATAGAAATTCACACAAATGACCTAAATATTTCTATGCTTTTGC 660
DB 635 TGGCTTAGCTGAAATCATAGAAATTCACACAAATGACCTAAATATTTCTATGCTTTTGC 694
QY 661 TTGTAAAGTTTCAGGACATGGAGGTGATAAAAAAATCTTTCTTAGGACAAATATGTAAAA 720
DB 695 TTGTAAAGTTTCAGGACATGGAGGTGATAAAAAAATCTTTCTTAGGACAAATATGTAAAA 754
QY 721 TGAATAAATTTCTAATCCCTGACTAATCTGAATGGAACCTCTTCTTAGGCCAAAGAGA 780
DB 755 TGAATAAATTTCTAATCCCTGACTAATCTGAATGGAACCTCTTCTTAGGCCAAAGAGA 814
QY 781 CCTCAGATGAACCTGAAGACCTGAATCTGGCCATGATAGGAGGAGGTGACACACACC 840
DB 815 CCTCAGATGAACCTGAAGACCTGAATCTGGCCATGATAGGAGGAGGTGACACACACC 874
QY 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGACAAAGTACCAGGATGAGTCAAGACT 900
DB 875 TTGTTATACCCCTTCCCTTTTGGAGTTTATGACAAAGTACCAGGATGAGTCAAGACT 934
QY 901 GATGAAATAGACTGATTTGGCAATTAAGAGTCCCAATTCACCTGACTCTGGTGTAGAT 960
DB 935 GATGAAATAGACTGATTTGGCAATTAAGAGTCCCAATTCACCTGACTCTGGTGTAGAT 994
QY 961 CACACACTGTCTGAGGATTTCCATCTATGAGACTTTTGTCTACATAACAGAGACTTTGGTT 1020
DB 995 CACACACTGTCTGAGGATTTCCATCTATGAGACTTTTGTCTACATAACAGAGACTTTGGTT 1054
QY 1021 TCACAACCCCTTTATTTTAGCTAAAGCAATCTTTTCTACTGACTTTCTTAAGTCTTTAGA 1080
DB 1055 TCACAACCCCTTTATTTTAGCTAAAGCAATCTTTTCTACTGACTTTCTTAAGTCTTTAGA 1114
QY 1081 CAAAGCTTAACCTTTCAACCAATTCGCAATCAGACAACTTTGAATCTACCTATGACCT 1140
DB 1115 CAAAGCTTAACCTTTCAACCAATTCGCAATCAGACAACTTTGAATCTACCTATGACCT 1174
QY 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGAAGTGAACCCATGTGCACTTTTCCA 1200
DB 1175 GTAAGCTCTCTCTGCTTCAAGATCTTGAAGTGAACCCATGTGCACTTTTCCA 1234
QY 1201 TTTAATGATTTATGCTTTGCTTTGAACTCTCTCTCCCTTAAATGTATAAAAAGTAAAG 1260
DB 1235 TTTAATGATTTATGCTTTGCTTTGAACTCTCTCTCCCTTAAATGTATAAAAAGTAAAG 1294
QY 1261 GTGACCTGACACCTCAGGCACACTTTCTCAGACCTCTCTGAGAGTGTATCCAGGCCAT 1320
DB 1295 GTGACCTGACACCTCAGGCACACTTTCTCAGACCTCTCTGAGAGTGTATCCAGGCCAT 1354
QY 1321 GGTAAAGTATGCTTTGCTTTGAACTCTCTCTCCCTTAAATGTATAAAAAGTAAAG 1380
DB 1355 GGTAAAGTATGCTTTGCTTTGAACTCTCTCTCCCTTAAATGTATAAAAAGTAAAG 1414
QY 1381 TTAACCAATAGCTTCCACAAATATATGTCCAAAGTCTTCAATTCGAGCTGTCAACCA 1440
DB 1415 TTAACCAATAGCTTCCACAAATATATGTCCAAAGTCTTCAATTCGAGCTGTCAACCA 1474
QY 1441 AATTTCAATGCAACATCTCCCATCCCAATTAACCTTTTCACTTTTGAAGTGTAACTA 1500
DB 1475 AATTTCAATGCAACATCTCCCATCCCAATTAACCTTTTCACTTTTGAAGTGTAACTA 1534
QY 1501 CTCATTAATAGTGTGAAGACCACTGACACCTCTTGTCTTAACTGACATTTTCACTTCAAT 1560
DB 1535 CTCATTAATAGTGTGAAGACCACTGACACCTCTTGTCTTAACTGACATTTTCACTTCAAT 1594
QY 1561 TTTTCTTTTCTATGATGATATTTTGTGATATATATATATATATATATATATATATAT 1620
DB 1595 TTTTCTTTTCTATGATGATATTTTGTGATATATATATATATATATATATATATATAT 1654
QY 1621 TAATAGTTTGTGACTTTGCTTTTCTGAGAGGAAATTTGAAAGTGTCAACAAAT 1675
DB 1655 TAATAGTTTGTGACTTTGCTTTTCTGAGAGGAAATTTGAAAGTGTCAACAAAT 1709

RESULT	9
AKI23798	
LOCUS	1709 bp mRNA linear PRI 09-SEP-2003
DEFINITION	Homo sapiens CDNA FLJ41804 fis, clone NOVAR2000710.
ACCESSION	AKI23798
VERSION	AKI23798.1 GI:34529425
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1
AUTHORS	Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T., Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R., Yoshioka,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiyama,H., Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E., Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S., Terashima,Y., Watanabe,M., Sugiyama,T., Irie,R., Otsuki,T., Sato.H., Wakamatsu,A., Ishii,S., Yamamoto,J., Iseno,Y., Kawai-Hto,Y., Saito,K., Nishikawa,T., Kimura,K., Yanashita,H., Matsum,C., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuna,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A..

QY 721 TGAATAAATTTCTAATCCCTGACTAACTGAATGAGCCCTCTCTTAGGCCAAGAGA 780
 Db 744 TGAATAAATTTCTAATCCCTGACTAACTGAATGAGCCCTCTCTTAGGCCAAGAGA 803
 QY 781 CTTAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAGGAGGTGAGACACCC 840
 Db 804 CTTAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGGAGGAGGTGAGACACCC 863
 QY 841 TTGTTATACCCCTTCCCTTTGGAGTTTATGACAAGTGCACAGGATGAGTCAATAGACT 900
 Db 864 TTGTTATACCCCTTCCCTTTGGAGTTTATGACAAGTGCACAGGATGAGTCAATAGACT 923
 QY 901 GATGAATAGACTGATTTGGCAATAGAGTCCCAATTCCAACCTGACTCTGCTGTAGAT 960
 Db 924 GATGAATAGACTGATTTGGCAATAGAGTCCCAATTCCAACCTGACTCTGCTGTAGAT 983
 QY 961 CACACACTGTCTGAGGGATTCATCTATGAGACTTTTGTCTACATAACAGAGACCTTGGTT 1020
 Db 984 CACACACTGTCTGAGGGATTCATCTATGAGACTTTTGTCTACATAACAGAGACCTTGGTT 1043
 QY 1021 TCCAGAACCCCTTTATTTAGCTTAAGACTTTCTTTCTACTGACTTTCTTAAGTCTTTAGA 1080
 Db 1044 TCCAGAACCCCTTTATTTAGCTTAAGACTTTCTTTCTACTGACTTTCTTAAGTCTTTAGA 1103
 QY 1081 CAAAGCTTAACCTTTCAACCAATTCGCAATCAGACAACTTTGAATCTACCTATGACCT 1140
 Db 1104 CAAAGCTTAACCTTTCAACCAATTCGCAATCAGACAACTTTGAATCTACCTATGACCT 1163
 QY 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGGCTCTTTAAGCTGAAACCGATGTGCACTTTCCA 1200
 Db 1164 GTAAGCTCTCTCTGCTTCAAGATCTTGGCTCTTTAAGCTGAAACCGATGTGCACTTTCCA 1223
 QY 1201 TTTAATGATTTATGCTTTGCTTGTAACTCTGCTCTCCCTAAATATGTAATAAGTAAAG 1260
 Db 1224 TTTAATGATTTATGCTTTGCTTGTAACTCTGCTCTCCCTAAATATGTAATAAGTAAAG 1283
 QY 1261 GTGACCTGACCACTCAGGCACACTTTCTCAGACCTCTCTGAGAGTGTATCCCAAGGCCAT 1320
 Db 1284 GTGACCTGACCACTCAGGCACACTTTCTCAGACCTCTCTGAGAGTGTATCCCAAGGCCAT 1343
 QY 1321 GGTAACTCATGTTGGCTCAGATCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTGG 1380
 Db 1344 GGTAACTCATGTTGGCTCAGATCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTGG 1403
 QY 1381 TTACCAATAGCTTCCACAAATATATGTCAGAACTCTCAATTCAGGCTCTGTCACCA 1440
 Db 1404 TTACCAATAGCTTCCACAAATATATGTCAGAACTCTCAATTCAGGCTCTGTCACCA 1463
 QY 1441 AATTTCAAAATGCCAAATCTCCCAATCCCAATTAACCTATTTTCACTTTGAGGTGTAATCTA 1500
 Db 1464 AATTTCAAAATGCCAAATCTCCCAATCCCAATTAACCTATTTTCACTTTGAGGTGTAATCTA 1523
 QY 1501 CTCAAATAAATGTTGTAAGACAGTACAGACCTTTTGTCTCAATCTGACATTTACTTCAAT 1560
 Db 1524 CTCAAATAAATGTTGTAAGACAGTACAGACCTTTTGTCTCAATCTGACATTTACTTCAAT 1583
 QY 1561 TTTTCTTTTCTATGTTACTGATATTTTGCATATAAATCTGAGTAATAGTTCAAAAT 1620
 Db 1584 TTTTCTTTTCTATGTTACTGATATTTTGCATATAAATCTGAGTAATAGTTCAAAAT 1643
 QY 1621 TAATAGTTTTGACATTTGGCTTTCTGAGAAGAAATTCGAAAGTGTCAAAAATAAAAA 1680
 Db 1644 TAATAGTTTTGACATTTGGCTTTCTGAGAAGAAATTCGAAAGTGTCAAAAATAAAAA 1703
 QY 1681 AGATGAATGAAGCATATATATGTCAAATTTTCAATTTTCTAGTCAACAGAGATC 1740
 Db 1704 AGATGAATGAAGCATATATATGTCAAATTTTCAATTTTCTAGTCAACAGAGATC 1763
 QY 1741 GAAGGATTTCTGTTCAATATAGTAAATAATGAAATATAGTGTGCTTATATTTGTTT 1800
 Db 1764 GAAGGATTTCTGTTCAATATAGTAAATAATGAAATATAGTGTGCTTATATTTGTTT 1823
 QY 1801 GCAACACACTAGTTTAACTTAACTGTGACTAGTTATCTCTACCGAAGGTGATGTAGT 1860

Db 1824 GCACACACTAGTTAATTTAACTGTGACTAGTTATCTCTACCGAAGGTGATGTAGT 1883
 QY 1861 TTCTGGTTTTAAAAATCAAGCAAACTGGAATAATCAATCAATTAATGCTTTCTTTCC 1920
 Db 1884 TTCTGGTTTTAAAAATCAAGCAAACTGGAATAATCAATCAATTAATGCTTTCTTTCC 1943
 QY 1921 AGAAGTTTTTAAATGATATGCCAGCTTCTAATTTTGGAGACAAAAGCCTTAATTCACAA 1980
 Db 1944 AGAAGTTTTTAAATGATATGCCAGCTTCTAATTTTGGAGACAAAAGCCTTAATTCACAA 2003
 QY 1981 TGCATTCAATATATTTTTTTTGTATAGTTACAGTATACAGTTAGTATCCCTTAGATG 2040
 Db 2004 TGCATTCAATATATTTTTTTTGTATAGTTACAGTATACAGTTAGTATCCCTTAGATG 2063
 QY 2041 AGATGCTTGGACACAGAGTGTTTTGGATTTTCAATTTTGTGATTTTGTGATTTTGAATATTT 2100
 Db 2064 AGATGCTTGGACACAGAGTGTTTTGGATTTTCAATTTTGTGATTTTGTGATTTTGAATATTT 2123
 QY 2101 CCATACATATATAGAGAGTTTGGAAATGGGATTCAGTCTAATCATAAATTTCACTTA 2160
 Db 2124 CCATACATATATAGAGAGTTTGGAAATGGGATTCAGTCTAATCATAAATTTCACTTA 2183
 QY 2161 TGTTCATATACACCTTATCTGAATAGCCTGAAGGTAATTTTATACAATTTTAAATAA 2220
 Db 2184 TGTTCATATACACCTTATCTGAATAGCCTGAAGGTAATTTTATACAATTTTAAATAA 2243
 QY 2221 TTTTATGCTGAAACAGAGTGTTCGCACATTTGGACCATCAGAAAGCAGAGTGTCACTAT 2280
 Db 2244 TTTTATGCTGAAACAGAGTGTTCGCACATTTGGACCATCAGAAAGCAGAGTGTCACTAT 2303
 QY 2281 TTCAGTCTAGTCTCAAAAAGTTTCAAGATGTTTAAAGTGTGATGTCAGTTTCATCCAGTGA 2340
 Db 2304 TTCAGTCTAGTCTCAAAAAGTTTCAAGATGTTTAAAGTGTGATGTCAGTTTCATCCAGTGA 2363
 QY 2341 TCCGAGTACTTTGGGAAGCCCAAGACAGGTGGATCTCTTGAGCCCGAGGTTTGAGGCCAG 2400
 Db 2364 TCCGAGTACTTTGGGAAGCCCAAGACAGGTGGATCTCTTGAGCCCGAGGTTTGAGGCCAG 2423
 QY 2401 ACTGCAACACAGTACAGACCTGTTTCTACAATAATTAATAAATTTAGCCAGTGTGGT 2460
 Db 2424 ACTGCAACACAGTACAGACCTGTTTCTACAATAATTAATAAATTTAGCCAGTGTGGT 2483
 QY 2461 GGTGCACACTGTGAGTCCAGGCTACTCAGAGGCTCAGGCTAGTAGTATGTTTGGAGACTG 2520
 Db 2484 GGTGCACACTGTGAGTCCAGGCTACTCAGAGGCTCAGGCTAGTAGTATGTTTGGAGACTG 2543
 QY 2521 GGAAGTTGAGGCTGAACTGAGCCAGGATCTTGCCACCATTCAGCTTGGGCAACAGAG 2580
 Db 2544 GGAAGTTGAGGCTGAACTGAGCCAGGATCTTGCCACCATTCAGCTTGGGCAACAGAG 2603
 QY 2581 TGAGACCTGTCTCAAAAAAATAAAAA 2608
 Db 2604 TGAGACCTGTCTCAAAAAAATAAAAA 2631

RESULT 8

LOCUS CQ842314 1709 bp DNA linear PAT 02-AUG-2004
 DEFINITION Sequence 961 from Patent EP1440981.
 ACCESSION CQ842314
 VERSION CQ842314.1 GI:50894101
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
 Yamanoto, J., Isono, Y., Nagai, K. and Irie, R.
 TITLE Full-length human cdna
 JOURNAL Patent: EP 1440981-A 961 28-JUL-2004;
 Research Association for Biotechnology (JP)

```

|||||
Db 2478 CCACAGTCTCTCTTTCCCTGAGTTTCCACGATCTCGCAACCAATGCTTGGTAT 2419
|||||
Qy 3473 AATTATAACTGTTCTAGATATTTGTAGCAATGTACCCCTTCCATATTTATTTTGTGTG 3532
|||||
Db 2418 AATTATAACTGTTCTAGATATTTGTAGCAATGTACCCCTTCCATATTTATTTTGTGTG 2359
|||||
Qy 3533 TAAGGCTTCTTTTAGTCATTAATAATTTTGTAGATTCATCATCTGTTTAATGTTCTATCA 3592
|||||
Db 2358 TAAGGCTTCTTTTAGTCATTAATAATTTTGTAGATTCATCATCTGTTTAATGTTCTATCA 2299
|||||
Qy 3593 GTAGTTGTACATCTTACTTGTCTGAGCATATCACCATATAGATATCTATATTTGTTAA 3652
|||||
Db 2298 GTAGTTGTACATCTTACTTGTCTGAGCATATCACCATATAGATATCTATATTTGTTAA 2239
|||||
Qy 3653 TCTAATCACTGATGATATGAGGATATTTAAGTTTGTGACATATGATTAAGTGGCTA 3712
|||||
Db 2238 TCTAATCACTGATGATATGAGGATATTTAAGTTTGTGACATATGATTAAGTGGCTA 2179
|||||
Qy 3713 TAAATGAATAAGGCACTACAAA 2156
|||||
Db 2178 TAAATGAATAAGGCACTACAAA 2156

```

```

RESULT 7
LOCUS BC024174 2648 bp mRNA linear PRI 01-APR-2004
DEFINITION Homo sapiens LOC148823, mRNA (cdna clone IMAGE:4109064).
ACCESSION BC024174
VERSION BC024174.1 GI:18848218
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schneitz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raja, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnier, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2648)

```

```

REFERENCE
AUTHORS Strausberg, R.
Direct Submission
Submitted (19-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

TITLE
JOURNAL NIH-MGC Project URL: http://mgc.nci.nih.gov
PUBMED Contact: MGC help desk
12477932 Email: gcgpc@nih.gov

```

```

REFERENCE
AUTHORS Tissue Procurement: ATCC
CNA Library Preparation: CLONTECH Laboratories, Inc.
CNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org

```

```

REMARK
COMMENT

```

```

contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 32 Row: b Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21687148.
FEATURES
Location/Qualifiers
1..2648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4109064"
/tissue_type="Bone marrow, acute myelogenous leukemia"
/clone_lib="NIH_MGC_55"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

```

```

source
1..2648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4109064"
/tissue_type="Bone marrow, acute myelogenous leukemia"
/clone_lib="NIH_MGC_55"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

```

```

ORIGIN
Query Match 69.2%; Score 2604.8; DB 9; Length 2648;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2606; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 GAGAAACCGAGTCTCTGTGAAAGATGGGAAATTTATCTCTGCGAAACCTCAGTTGCCTG 60
Db 24 GAGAAACCGAGTCTCTGTGAAAGATGGGAAATTTATCTCTGCGAAACCTCAGTTGCCTG 83
Qy 61 GGAGAGAAATCAAAAGAGCCCAAGAGGAAACCCAGATGAGGAGAGAGAGAGAGAGAGAG 120
Db 84 GGAGAGAAATCAAAAGAGCCCAAGAGGAAACCCAGATGAGGAGAGAGAGAGAGAGAGAG 143
Qy 121 ATGACTACATTTGAAAGAGAAATTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAG 180
Db 144 ATGACTACATTTGAAAGAGAAATTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAG 203
Qy 181 ACTTTAATCAGGAAACGAGAAATGGCAGTGGTCTGAGAAAGTGTGCTACATGTCATT 240
Db 204 ACTTTAATCAGGAAACGAGAAATGGCAGTGGTCTGAGAAAGTGTGCTACATGTCATT 263
Qy 241 AATCAGATCCCATCAGAGATCTCCCTGAGTCCCAATGATGATGATGATGATGATGATGAT 300
Db 264 AATCAGATCCCATCAGAGATCTCCCTGAGTCCCAATGATGATGATGATGATGATGATGAT 323
Qy 301 GACTCCCTCAAGAGAAAGTGAGACAGTTTATAGAGAAAGGTGAGAGAGAGAGAGAGAGAG 360
Db 324 GACTCCCTCAAGAGAAAGTGAGACAGTTTATAGAGAAAGGTGAGAGAGAGAGAGAGAGAG 383
Qy 361 CTTAGGACTTCTGTTAGTAGGCTTGTCTCTGACCCATCAGCATGATGATGATGATGATGAT 420
Db 384 CTTAGGACTTCTGTTAGTAGGCTTGTCTCTGACCCATCAGCATGATGATGATGATGATGAT 443
Qy 421 TTTCCACACTAAATCTCAAGCTGCTTTATACCTTCCAGCAATGAAGCAATGAGAGAGAG 480
Db 444 TTTCCACACTAAATCTCAAGCTGCTTTATACCTTCCAGCAATGAAGCAATGAGAGAGAG 503
Qy 481 TAGCAGACTCTGCGAAAGTTGTTTACCCTGAGCAGTGCATGAAACATTCCTTTCTGGCTA 540
Db 504 TAGCAGACTCTGCGAAAGTTGTTTACCCTGAGCAGTGCATGAAACATTCCTTTCTGGCTA 563
Qy 541 AAGTTTAGAAATATATCTTTATATATATATATATATATATATATATATATATATATAT 600
Db 564 AAGTTTAGAAATATATCTTTATATATATATATATATATATATATATATATATATATAT 623
Qy 601 TGCTTTAGGTGAATCATAGAAATTCACACAATGACCTAAATATTTCTATGCTTTTTCG 660
Db 624 TGCTTTAGGTGAATCATATATATATATATATATATATATATATATATATATATATAT 683
Qy 661 TTGTTAAAGTTTGAGGACATGGAGGTGATATAAAAAAACTTTCTTAGGACAAATATGTA 720
Db 684 TTGTTAAAGTTTGAGGACATGGAGGTGATATAAAAAAACTTTCTTAGGACAAATATGTA 743

```

QY 1253 AGTAAACGGTGACCTGACCACTCAGGACACCTTCTCAGGACCTCCTCAGAGTGTATCC 1312
 Db 4637 AGTAAACGGTGACCTGACCACTCAGGACACCTTCTCAGGACCTCCTCAGAGTGTATCC 4578
 QY 1313 CAGGCCATCGTAAAGTCATGCTGGCTCAGAAATCAACCTCTTTAAATATTTTACAGAAATTG 1372
 Db 4577 CAGGCCATCGTAAAGTCATGCTGGCTCAGAAATCAACCTCTTTAAATATTTTACAGAAATTG 4518
 QY 1373 GGTGTTGGTTACCAATTAAGTCTCCACAAATATATGTCGAAGAACTTCAATTTCCAGCCT 1432
 Db 4517 GGTGTTGGTTACCAATTAAGTCTCCACAAATATATGTCGAAGAACTTCAATTTCCAGCCT 4458
 QY 1433 GCTCACCAAATTTCAAAATGCCAACATCTCCCAATCAATTAAGTCTTTTCAATTTTCAAGGT 1492
 Db 4457 GCTCACCAAATTTCAAAATGCCAACATCTCCCAATCAATTAAGTCTTTTCAATTTTCAAGGT 4398
 QY 1493 GTAATCTCTCAATAAATCTGTGAAGCAGTGACAGACCCCTTTCTAACTGACATTT 1552
 Db 4397 GTAATCTCTCAATAAATCTGTGAAGCAGTGACAGACCCCTTTCTAACTGACATTT 4338
 QY 1553 ACTTCAATTTTCTTTTCTATGTAAGTATTTTTGATATAAATCTTGCAAGTAAATAGT 1612
 Db 4337 ACTTCAATTTTCTTTTCTATGTAAGTATTTTTGATATAAATCTTGCAAGTAAATAGT 4278
 QY 1613 TCAAAAATTAATAGTTTTGACANTGGCTTTCTGAGAGAGAAATTTGAAAGTGCACAA 1672
 Db 4277 TCAAAAATTAATAGTTTTGACANTGGCTTTCTGAGAGAGAAATTTGAAAGTGCACAA 4218
 QY 1673 AATAAAAAGATGAATGAAGCATATATAAATGTCAATTTTTCAATTTTCTAGTCAAC 1732
 Db 4217 AATAAAAAGATGAATGAAGCATATATAAATGTCAATTTTTCAATTTTCTAGTCAAC 4158
 QY 1733 AGAGAAATCGAAGATTCGTTCAAATATATAGTAAATAATGAAATTAATCTTGTCATATA 1792
 Db 4157 AGAGAAATCGAAGATTCGTTCAAATATATAGTAAATAATGAAATTAATCTTGTCATATA 4098
 QY 1793 TTTTGTTCGACACATAGTTAATTAACCTGTGACTAGTATCTACCGAGGTGGA 1852
 Db 4097 TTTTGTTCGACACATAGTTAATTAACCTGTGACTAGTATCTACCGAGGTGGA 4038
 QY 1853 TGTGATGTTCTCGTTTAAATTTCAAGCAAACTCGAAATAATCCATCTAATATGCTT 1912
 Db 4037 TGTGATGTTCTCGTTTAAATTTCAAGCAAACTCGAAATAATCCATCTAATATGCTT 3978
 QY 1913 TCTTTCCCAAGAGTTTTTAATGATATGCGAGCTTCTTAATTTGGAGACAAAGCCTTA 1972
 Db 3977 TCTTTCCCAAGAGTTTTTAATGATATGCGAGCTTCTTAATTTGGAGACAAAGCCTTA 3918
 QY 1973 ATTGCAATGCAATTCATATATATTTTTTGTATAGTTACAGTATACGAGTTCAGTATCC 2032
 Db 3917 ATTGCAATGCAATTCATATATATTTTTTGTATAGTTACAGTATACGAGTTCAGTATCC 3858
 QY 2033 CTTAGATGAGATGCTTTGGACCAAGAGTGTTTTGGATTTTCAAGTATTTTGGATTTTG 2092
 Db 3857 CTTAGATGAGATGCTTTGGACCAAGAGTGTTTTGGATTTTCAAGTATTTTGGATTTTG 3798
 QY 2093 GAATATTTCCATACATATATATGAGAGATTGGAAATGGGATTCAGTCTAATCATATAA 2152
 Db 3797 GAATATTTCCATACATATATGAGAGATTGGAAATGGGATTCAGTCTAATCATATAA 3738
 QY 2153 TTCACATTTGTTGATATACACCTTATCTGAATAGCTGAAGGTAAATTTTATACAAATTT 2212
 Db 3737 TTCACATTTGTTGATATACACCTTATCTGAATAGCTGAAGGTAAATTTTATACAAATTT 3678
 QY 2213 TTAATAATTTTATGCTGAAACAGAGTTTTGCGACATTTGGACCATCAGAAAGCAGAGT 2272
 Db 3677 TTAATAATTTTATGCTGAAACAGAGTTTTGCGACATTTGGACCATCAGAAAGCAGAGT 3618
 QY 2273 GTCACTATTTCAAGTCAGTCTCAAAAAGTTTTGAGTGTAAAGTGGTGTAGTCAATTCAT 2332
 Db 3617 GTCACTATTTCAAGTCAGTCTCAAAAAGTTTTGAGTGTAAAGTGGTGTAGTCAATTCAT 3558

QY 2333 GCCAGTGATCCGAGTACTTTGGAAAGCCAAAGCAGGTGGATCTCTTGAGCCAGGAGTTT 2392
 Db 3557 GCCAGTGATCCGAGTACTTTGGAAAGCCAAAGCAGGTGGATCTCTTGAGCCAGGAGTTT 3498
 QY 2393 GAGGCCAGACTCCACAAACACAGTGAGACCTCGTTTCTACAAATAATTTAAAAAATTAGCCA 2452
 Db 3497 GAGGCCAGACTCCACAAACACAGTGAGACCTCGTTTCTACAAATAATTTAAAAAATTAGCCA 3438
 QY 2453 GGTGTGGTGGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGGATTTGTT 2512
 Db 3437 GGTGTGGTGGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGGATTTGTT 3378
 QY 2513 TGAGACTGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCACCAATTTCCAGCTTTGG 2572
 Db 3377 TGAGACTGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCACCAATTTCCAGCTTTGG 3318
 QY 2573 CAACAGAGTGAGACCTGTCTCAAAAAAAGTTTCAAGATTTTGGAGATTTTCAAG 2632
 Db 3317 CAACAGAGTGAGACCTGTCTCAAAAAAAGTTTCAAGATTTTGGAGATTTTCAAG 3259
 QY 2633 ATCTTCAGATTTAGGATTTTCAACCTGTACTGACCTTTTGTAGTCATTTGCAAGCATTAATC 2692
 Db 3258 ATCTTCAGATTTAGGATTTTCAACCTGTACTGACCTTTTGTAGTCATTTGCAAGCATTAATC 3199
 QY 2693 AATAGTGACTCCAGATAACTCATTTGCTGTATACACATTTTGCCTCTCTATTCAACGA 2752
 Db 3198 AATAGTGACTCCAGATAACTCATTTGCTGTATACACATTTTGCCTCTCTATTCAACGA 3139
 QY 2753 ATTCTTATCCCTCTTGTGTTAATGTCGGAAGGAAACAAATAGAAATTTTGGCA 2812
 Db 3138 ATTCTTATCCCTCTTGTGTTAATGTCGGAAGGAAACAAATAGAAATTTTGGCA 3079
 QY 2813 ATTCTAGAAAGTCATTTCTGCAAAATATGTCAGTCTCTGTAGATATTAGCAATTTTGG 2872
 Db 3078 ATTCTAGAAAGTCATTTCTGCAAAATATGTCAGTCTCTGTAGATATTAGCAATTTTGG 3019
 QY 2873 AAAAAACAAAAATTTTTTACTTTTCTGCTGCTTTGCTAGCTGTTTATGATATAAATPACC 2932
 Db 3018 AAAAAACAAAAATTTTTTACTTTTCTGCTGCTTTGCTAGCTGTTTATGATATAAATPACC 2959
 QY 2933 TTAATTTGTAATAAATAATTTTAAATTTGAGTAAACAAATCTGGAATTTATCAGAAAGGGC 2992
 Db 2958 TTAATTTGTAATAAATAATTTTAAATTTGAGTAAACAAATCTGGAATTTATCAGAAAGGGC 2899
 QY 2993 AAGCAATAGTTAATAAAGATTTGATTTGTTAGAGGACGTTGAAATCCAGAGCATC 3052
 Db 2898 AAGCAATAGTTAATAAAGATTTGATTTGTTAGAGGACGTTGAAATCCAGAGCATC 2839
 QY 3053 AATGCTCTTCTGTTGTTTCAACCAAGCCAGAGATGCTTAATCTTTCCGAGATCTAG 3112
 Db 2838 AATGCTCTTCTGTTGTTTCAACCAAGCCAGAGATGCTTAATCTTTCCGAGATCTAG 2779
 QY 3113 TTTTTCAGCAAGCAGGATTTAAGAAATGTAATCTTATGTTGGTTATGAGAAACAAATA 3172
 Db 2778 TTTTTCAGCAAGCAGGATTTAAGAAATGTAATCTTATGTTGGTTATGAGAAACAAATA 2719
 QY 3173 GAATCATCTGTTATAGTCTTTTAACTGTAATAATTTTGTGAAGCTTATCTTTTATGC 3232
 Db 2718 GAATCATCTGTTATAGTCTTTTAACTGTAATAATTTTGTGAAGCTTATCTTTTATGC 2659
 QY 3233 ATATAAATAATTTGAACATTTTATATTTTTTAACTAGTTTTTACTCAAGTGTGA 3292
 Db 2658 ATATAAATAATTTGAACATTTTATATTTTTTAACTAGTTTTTACTCAAGTGTGA 2599
 QY 3293 TTAATATCAAGAAAAATGTAACCACTGTAAGGTTAGAGTTATAGAAATTTTGTCAATGTA 3352
 Db 2598 TTAATATCAAGAAAAATGTAACCACTGTAAGGTTAGAGTTATAGAAATTTTGTCAATGTA 2539
 QY 3353 TTAACCCATGTTAGTCACTCTTTATGAAGAGACAGAACCTTACATCTCCCAAGAAAGTT 3412
 Db 2538 TTAACCCATGTTAGTCACTCTTTATGAAGAGACAGAACCTTACATCTCCCAAGAAAGTT 2479
 QY 3413 CCACAGTGCTCTTTTCCCTGAGTTTCAACAGTCTGCGCAACCAATGATCTGCTTCGTAT 3472

Sequencing vector: M13; 100%
 Sequencing vector: plasmid; 0%
 Chemistry: Dye-terminator ET; 100% of reads
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 136274 bases at least Q40
 Consensus quality: 138356 bases at least Q30
 Consensus quality: 139182 bases at least Q20
 Insert size: 142000; agarose-fp
 Insert size: 140368; sum-of-contigs
 Quality coverage: 4.64 in Q20 bases; agarose-fp
 Quality coverage: 4.81 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 10 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 31377: contig of 31377 bp in length
 31378 31477: gap of unknown length
 31478 31478: contig of 31534 bp in length
 63011: gap of unknown length
 63012 63111: gap of unknown length
 63112 64845: contig of 1734 bp in length
 64846 64945: gap of unknown length
 64946 67606: contig of 2661 bp in length
 67607 67706: gap of unknown length
 67707 77674: contig of 9968 bp in length
 77675 77774: gap of unknown length
 77775 89256: contig of 11482 bp in length
 89257 89357: gap of unknown length
 89357 100400: contig of 11044 bp in length
 100401 100500: gap of unknown length
 100501 111915: contig of 11415 bp in length
 111916 112015: gap of unknown length
 112016 125713: contig of 13698 bp in length
 125714 125814: gap of unknown length
 125814 141268: contig of 15455 bp in length.

FEATURES

source

1. 141268
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-115C4"

misc_feature
 1. 31377
 /note="assembly_name:Contig10"
 misc_feature
 31478..63011
 /note="assembly_name:Contig11"
 misc_feature
 63112..64845
 /note="assembly_name:Contig2"
 misc_feature
 64946..67606
 /note="assembly_name:Contig3"
 misc_feature
 67707..77674
 /note="assembly_name:Contig4"
 misc_feature
 77775..89256
 /note="assembly_name:Contig5"
 misc_feature
 89357..100400
 /note="assembly_name:Contig6"
 misc_feature
 100501..111915
 /note="assembly_name:Contig7"
 misc_feature
 112016..125713
 /note="assembly_name:Contig8"
 misc_feature
 125814..141268
 /note="assembly_name:Contig9"
 clone_end:SP6
 vector_side:left

ORIGIN

Query Match 93.3%; Score 3511.6; DB 2; Length 141268;
 Best Local Similarity 99.3%; Pred. No. 0;

	Matches 3537;	Conservative	0;	Mismatches	25;	Indels	1;	Gaps	1;
Qy	173	TTTTCATCCACTTCTTAATCAGGAAAAACGAGAAATGGCAGTGGTCTCTGAAGAAAGTGTGTACA	232						
Db	5717	TTTCTCTTTTGTGTTGGCAGGAAAAACGAGAAATGGCAGTGGTCTCTGAAGAAAGTGTGTACA	5658						
Qy	233	CTGTCAATTAATCACATCCCAATCAGAGATCCCTCCCTGAGCTCCAATGATGATGGCTATG	292						
Db	5657	CTGTCAATTAATCACATCCCAATCAGAGATCCCTCCCTGAGCTCCAATGATGATGGCTATG	5598						
Qy	293	AGAACATTGACTCCCTCACAGGAAAGTGAGACAGTTTATAGAAAGTCTAGAGACACAAT	352						
Db	5597	AGAACATTGACTCCCTCACAGGAAAGTGAGACAGTTTATAGAAAGTCTAGAGACACAAT	5538						
Qy	353	ATGCCCTTCTTAGGACTTCTGTCTAGTAGGCTTGTCTCTGCACCCATGAGCATGATTATG	412						
Db	5537	ATGCCCTTCTTAGGACTTCTGTCTAGTAGGCTTGTCTCTGCACCCATGAGCATGATTATG	5478						
Qy	413	AAGTTGTGTTTCCACACTAAATCCCTCAAGCTGTTTATACCTTCCAGCAATGAAGACA	472						
Db	5477	AAGTTGTGTTTCCACACTAAATCCCTCAAGCTGTTTATACCTTCCAGCAATGAAGACA	5418						
Qy	473	ATGCAGAAATAGCAGACTCTCGCGAAGTTGTTCACCTGACGATGCAATGAACATTCCTT	532						
Db	5417	ATGCAGAAATAGCAGACTCTCGCGAAGTTGTTCACCTGACGATGCAATGAACATTCCTT	5358						
Qy	533	TCTGGCTAAAGTTTAGAAATATTATCTTATTATATATATATATATATATATATATATAT	592						
Db	5357	TCTGGCTAAAGTTTAGAAATATTATCTTATTATATATATATATATATATATATATATAT	5298						
Qy	593	CATCTCTGTGGCTTAGTGGAATCATAGAAATGACACAATGACCTAATAATTTCTATGT	652						
Db	5297	CATCTCTGTGGCTTAGTGGAATCATAGAAATGACACAATGACCTAATAATTTCTATGT	5238						
Qy	653	GTTTTGTCTGTAAGTTTCAGGACATGAGAGTGTATATAAATAAATTTCTTAGGACAATA	712						
Db	5237	GTTTTGTCTGTAAGTTTCAGGACATGAGAGTGTATATAAATAAATTTCTTAGGACAATA	5178						
Qy	713	ATGTAAATGAAATAAATTTCTAATCCCTCCCTGACTAATCTGAATGAGCCCTCTTCTAGGC	772						
Db	5177	ATGTAAATGAAATAAATTTCTAATCCCTCCCTGACTAATCTGAATGAGCCCTCTTCTAGGC	5118						
Qy	773	CAAGAGACCTCAGATGAACCTGAAGACTGAATTTCTGGCCATGATAGGAGGAGGTGA	832						
Db	5117	CAAGAGACCTCAGATGAACCTGAAGACTGAATTTCTGGCCATGATAGGAGGAGGTGA	5058						
Qy	833	GACACACCTTGTATACCCCTTCCCTTTTGGAGTTTATGCACAAGTGACACGAGATGATC	892						
Db	5057	GACACACCTTGTATACCCCTTCCCTTTTGGAGTTTATGCACAAGTGACACGAGATGATC	4998						
Qy	893	ATAAGACTGATGAAATAGACTGATTGTGGCAATAAGAGTCCCAATTCACACCTGACTCTG	952						
Db	4997	ATAAGACTGATGAAATAGACTGATTGTGGCAATAAGAGTCCCAATTCACACCTGACTCTG	4938						
Qy	953	GTGTAGATCACACTGTCTGAGGATTCATCTATGAGACTTTGTCTACATACACAGAGA	1012						
Db	4937	GTGTAGATCACACTGTCTGAGGATTCATCTATGAGACTTTGTCTACATACACAGAGA	4878						
Qy	1013	CCTTGGTTTCCACACACCCCTTTATTTTAGCTAAAGCATTTCTTTCTACTGACTTCTTAAG	1072						
Db	4877	CCTTGGTTTCCACACACCCCTTTATTTAGCTAAAGCATTTCTTTCTACTGACTTCTTAAG	4818						
Qy	1073	TCTTTAGACAAAGCTTAATCTTTTCAACCAATGCGCAATTCAGACAACTTTGAATCTACC	1132						
Db	4817	TCTTTAGACAAAGCTTAATCTTTTCAACCAATGCGCAATTCAGACAACTTTGAATCTACC	4758						
Qy	1133	TATGACTGTAACTCTCTCCCTGCTTCAAGATCTTGCCTCTTTAAGCTGACACGATGTC	1192						
Db	4757	TATGACTGTAACTCTCTCCCTGCTTCAAGATCTTGCCTCTTTAAGCTGACACGATGTC	4698						
Qy	1193	ACTTTCCATTTAATGATTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	1252						
Db	4697	ACTTTCCATTTAATGATTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	4638						

Db	133923	CTTAGATGAGATGCTTGGGACCAAGATGTTTTGGATTTTCAGATTTATTTTGGATTTTG	133982
QY	2093	GAATATTTCCATACATATTAATGAGAGAGTTGAAATAATGGGATTCAGTCTTAATCATATAA	2152
Db	133983	GAATATTTCCATACATATTAATGAGAGAGTTGAAATAATGGGATTCAGTCTTAATCATATAA	134042
QY	2153	TTCACTTATGTTTGATATACACCTTATCTGAATAGCTGCAAGGTAAATTTTATACAAATTT	2212
Db	134043	TTCACTTATGTTTGATATACACCTTATCTGAATAGCTGCAAGGTAAATTTTATACAAATTT	134102
QY	2213	TTAAATAATTTTATGCTGAAACAGAGTTTGGCCACATTTGGACCATTCAGAAAGCAGAAGT	2272
Db	134103	TTAAATAATTTTATGCTGAAACAGAGTTTGGCCACATTTGGACCATTCAGAAAGCAGAAGT	134162
QY	2273	GTCACTATTTCAAGTCAGTCTCAAAAAGTTTTCAGATGTTAAAGTGTGATCGAGTTTCAAT	2332
Db	134163	GTCACTATTTCAAGTCAGTCTCAAAAAGTTTTCAGATGTTAAAGTGTGATCGAGTTTCAAT	134222
QY	2333	GCAGTGATCCGAGTACTTTGGGAAGCCCAAGACAGGTGGATCTCTTGAGCCCGAGGATTT	2392
Db	134223	GCAGTGATCCGAGTACTTTGGGAAGCCCAAGACAGGTGGATCTCTTGAGCCCGAGGATTT	134282
QY	2393	GAGGCCAGACTGCCAACAACAGTGAGACCTCGTTTCTACAAATAATTTAAAAAATTAGCCA	2452
Db	134283	GAGGCCAGACTGCCAACAACAGTGAGACCTCGTTTCTACAAATAATTTAAAAAATTAGCCA	134342
QY	2453	GGTGTGGTGGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGGATTTT	2512
Db	134343	GGTGTGGTGGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGGATTTT	134402
QY	2513	TCGAGCTGGGAGTTTGAGGCTGAACTGAGCCAGGATCTTGGCCACCACTTCAGAGCTTGGG	2572
Db	134403	TCGAGCTGGGAGTTTGAGGCTGAACTGAGCCAGGATCTTGGCCACCACTTCAGAGCTTGGG	134462
QY	2573	CAACGAGTGAGACCCCTGTCTCAAAAAAATAAAGTTTTCAGATTTTGGAGCATTTGAG	2632
Db	134463	CAACGAGTGAGACCCCTGTCTCAAAAAAATAAAGTTTTCAGATTTTGGAGCATTTGAG	134521
QY	2633	ATCTTCAGATTTAGGATTTTCAACCTGTACTGACCTTTTGTAGTCAATGACAGCATTAATC	2692
Db	134522	ATCTTCAGATTTAGGATTTTCAACCTGTACTGACCTTTTGTAGTCAATGACAGCATTAATC	134581
QY	2693	AATAGTGAGCTCCAGATAACTCATTTGCTGTATACACATTTTGGCTCTCTATTCAACGA	2752
Db	134582	AATAGTGAGCTCCAGATAACTCATTTGCTGTATACACATTTTGGCTCTCTATTCAACGA	134641
QY	2753	ATCTTATCCCTCTTGTGGTGAATTTAATGTGCGAAGGAAACAATAGAAATTTTGA	2812
Db	134642	ATCTTATCCCTCTTGTGGTGAATTTAATGTGCGAAGGAAACAATAGAAATTTTGA	134701
QY	2813	ATCTAGAAAGTCATTTCTGCAAAATATGTGAGTCTGATATATTAGCCAAATTTAGG	2872
Db	134702	ATCTAGAAAGTCATTTCTGCAAAATATGTGAGTCTGATATATTAGCCAAATTTAGG	134761
QY	2873	AAAAACACAAAAATTTTACTTTTCTGCTGCTTCTGCTGCTTTTATGATATAAATACC	2932
Db	134762	AAAAACACAAAAATTTTACTTTTCTGCTGCTTCTGCTGCTTTTATGATATAAATACC	134821
QY	2933	TTATTTGTAATAAATAATTTTAAATTTGAGTAAACAACTCGGAATATCAGAGAAGGGC	2992
Db	134822	TTATTTGTAATAAATAATTTTAAATTTGAGTAAACAACTCGGAATATCAGAGAAGGGC	134881
QY	2993	AGCAATAGTTTAAATAACAGATTTGATTTGGTAGAGGACGTTGAAATCCAGAGCATC	3052
Db	134882	AGCAATAGTTTAAATAACAGATTTGATTTGGTAGAGGACGTTGAAATCCAGAGCATC	134941
QY	3053	AATGCTTCTTGTGGTGTCCACATAAGCCACAGCAGATGTCTTAATCTTTCCGAGATCTAG	3112
Db	134942	AATGCTTCTTGTGGTGTCCACATAAGCCACAGCAGATGTCTTAATCTTTCCGAGATCTAG	135001
QY	3113	TTTTTCAGCAAGCAGAGTTTAAAGAAATGTAACTATCTATGTTGGTTATGAAGAAACAATA	3172
Db	135002	TTTTTCAGCAAGCAGAGTTTAAAGAAATGTAACTATCTATGTTGGTTATGAAGAAACAATA	135061

RESULT 6
 AC074365/c
 LOCUS 141268 bp DNA linear HTG 23-SEP-2000
 DEFINITION Homo sapiens chromosome 1 clone RP11-115C4, WORKING DRAFT SEQUENCE,
 10 unordered pieces.
 AC074365
 AC074365.5 GI:10280935
 VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Waterston,R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 141268)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (29-JUL-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 COMMENT On Sep 23, 2000 this sequence version replaced gi:9838075.
 ----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H_NH0115C04
 ----- Summary Statistics -----

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

FEATURES

Location/Qualifiers
1..185467
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-978115"
/clone_lib="RPC1-11.4"

ORIGIN

Query Match 93.6%; Score 3522.2; DB 9; Length 185467;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 3544; Conservative 0; Mismatches 18; Indels 1; Gaps 1;
QY 173 TTTCATCCACCTTCTAATCAGGAAACGAGAAATGGCAGTGGTCTGGAAGAGTGCTAC 232
Db 132063 TTTCCTTTGGCTTGGCAGGAAACGAGAAATGGCAGTGGTCTGGAAGAGTGCTAC 132122
QY 233 CTGTCATTAAATCACAATCCCATCAGAGATCTCCCTGAGCTCCAAATGATGCTATG 292
Db 132123 CTGTCATTAAATCACAATCCCATCAGAGATCTCCCTGAGCTCCAAATGATGCTATG 132182
QY 293 AGACATTGATCTCCCTCAAGAGAAAGTGACAGTCTTGTAGAAAGGTTCAGACAGAAAT 352
Db 132183 AGACATTGATCTCCCTCAAGAGAAAGTGACAGTCTTGTAGAAAGGTTCAGACAGAAAT 132242
QY 353 ATGCCCTCTTAGGACTCTCTGTAGTGGCTTGTCTGACCCATGAGCATGATG 412
Db 132243 ATGCCCTCTTAGGACTCTCTGTAGTGGCTTGTCTGACCCATGAGCATGATG 132302
QY 413 AAGTTGTGTTTCCACACTAAATCCTCAAGCTGCTTTATCACTTCCAGCAATCAAGACA 472
Db 132303 AAGTTGTGTTTCCACACTAAATCCTCAAGCTGCTTTATCACTTCCAGCAATCAAGACA 132362
QY 473 ATGCAGAAATAGCAGACTCTGGCGAAGTGTTCACCCGTGAGCAGTGCATGAACATTCCTT 532
Db 132363 ATGCAGAAATAGCAGACTCTGGCGAAGTGTTCACCCGTGAGCAGTGCATGAACATTCCTT 132422
QY 533 TCTGGCTAAAGTTTAGAAATATATCTATTATATATCCCTTAGCAACTCTGATATGG 592
Db 132423 TCTGGCTAAAGTTTAGAAATATATCTATTATATATCCCTTAGCAACTCTGATATGG 132482
QY 593 CATCTGTGGCTTAGGTGAATCATAGAAATGACACAAATGACCTAAATATTTCTATGT 652
Db 132483 CATCTGTGGCTTAGGTGAATCATAGAAATGACACAAATGACCTAAATATTTCTATGT 132542
QY 653 GTTTTCTGTTAAAGTTTAGGACATGGAGTGATATAAAAAAATTTCTTAGGACAATA 712
Db 132543 GTTTTCTGTTAAAGTTTAGGACATGGAGTGATATAAAAAAATTTCTTAGGACAATA 132602
QY 713 ATGTAAATGAAATAAATTTCTAATCCCTGACTAATCTGAATGGACCTCTCTAGGC 772
Db 132603 ATGTAAATGAAATAAATTTCTAATCCCTGACTAATCTGAATGGACCTCTCTAGGC 132662
QY 773 CAAAGAGACTCAGATGAACCTGAAGACTGAATTTCTGGCCATGATAGGAGGAGTGA 832
Db 132663 CAAAGAGACTCAGATGAACCTGAAGACTGAATTTCTGGCCATGATAGGAGGAGTGA 132722
QY 833 GACACACTGTTATACCCCTTCCCTTTTGGAGTTTATGACAAAGTACCAGGATGATC 892
Db 132723 GACACACTGTTATACCCCTTCCCTTTTGGAGTTTATGACAAAGTACCAGGATGATC 132782
QY 893 ATAAGACTGATGAATAGACTGATTGGCAATAAGAGTCCCAATTCACCTGACTCTG 952
Db 132783 ATAAGACTGATGAATAGACTGATTGGCAATAAGAGTCCCAATTCACCTGACTCTG 132842
QY 953 GTGTAGATCACACTGTCTGAGGGATTCCATCTATGAGACTTTGTCTACATAACAGAGA 1012

Db 132843 GTGTAGATCACACTGTCTGAGGATTCCTATCTATGAGACTTTGTCTACATAACAGAGA 132902
QY 1013 CCTTGGTTTCCACAACCCCTTATTTAGCTTAAAGCATTCTTTCTACTGACTTCTTAAG 1072
Db 132903 CCTTGGTTTCCACAACCCCTTATTTAGCTTAAAGCATTCTTTCTACTGACTTCTTAAG 132962
QY 1073 TCTTTAGCAAAAGCTTAACTCTTTCAACCAATTCGAATCAGACAAACTTTGAATCTACC 1132
Db 132963 TCTTTAGCAAAAGCTTAACTCTTTCAACCAATTCGAATCAGACAAACTTTGAATCTACC 133022
QY 1133 TATGACCTGTAAGCTCTCTCTGCTTCAAGATCTTGGCTCTTTAAGCTGAACCGATGTC 1192
Db 133023 TATGACCTGTAAGCTCTCTCTGCTTCAAGATCTTGGCTCTTTAAGCTGAACCGATGTC 133082
QY 1193 ACTTTCCATTTAATGATTTATGCTTTGCTTGTAACTCTCTGCTCTCCCTTAAATGTATAA 1252
Db 133083 ACTTTCCATTTAATGATTTATGCTTTGCTTGTAACTCTCTGCTCTCCCTTAAATGTATAA 133142
QY 1253 AGTAAACGGTGACCTGACCACTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCC 1312
Db 133143 AGTAAACGGTGACCTGACCACTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCC 133202
QY 1313 CAGCCCATGTAAGTCAATGTTGGCTCAGATCAACCTCTTTAAATATTTTACAGAAATTTG 1372
Db 133203 CAGCCCATGTAAGTCAATGTTGGCTCAGATCAACCTCTTTAAATATTTTACAGAAATTTG 133262
QY 1373 GGTTTTGGTTTACCATAAAGTCTCCACAAATATATGTCCAAAGATCTTCAATTCGAAGCT 1432
Db 133263 GGTTTTGGTTTACCATAAAGTCTCCACAAATATATGTCCAAAGATCTTCAATTCGAAGCT 133322
QY 1433 GCTCACCAATTTCAAATGCCAATCTCCCATCAATTTACCTATTTTCATCTTTGAGGT 1492
Db 133323 GCTCACCAATTTCAAATGCCAATCTCCCATCAATTTACCTATTTTCATCTTTGAGGT 133382
QY 1493 GTAACTCTACTCAATAAATCTGTGAAGCAGTGAAGCAGACCTTTGTCTAAGCTGACATTT 1552
Db 133383 GTAACTCTACTCAATAAATCTGTGAAGCAGTGAAGCAGACCTTTGTCTAAGCTGACATTT 133442
QY 1553 ACTTCAATTTTCTTTTCTATGCTACTGATATTTTGTGATATAAACTTCAGTCAATAGT 1612
Db 133443 ACTTCAATTTTCTTTTCTATGCTACTGATATTTTGTGATATAAACTTCAGTCAATAGT 133502
QY 1613 TCAAAATTAATAGTTTTCAGATTTGGCTTTCTGAGAAGAAATTTGAAAGTGTCAAA 1672
Db 133503 TCAAAATTAATAGTTTTCAGATTTGGCTTTCTGAGAAGAAATTTGAAAGTGTCAAA 133562
QY 1673 AATAAAAAAGATGAATGAAGCATATATAATTTGTCAATTTTTCATTTCTAGTCAAC 1732
Db 133563 AATAAAAAAGATGAATGAAGCATATATAATTTGTCAATTTTTCATTTCTAGTCAAC 133622
QY 1733 AGAATCGAAGATTTCTGTTCAATTTAGTAAAAATTAAGAAATTAAGTGTCTTATA 1792
Db 133623 AGAATCGAAGATTTCTGTTCAATTTAGTAAAAATTAAGAAATTAAGTGTCTTATA 133682
QY 1793 TTTTGTGTCACACACTAGTTTAACTTGAACCTGTGCTAGTTATCTCTACCGAAGTGA 1852
Db 133683 TTTTGTGTCACACACTAGTTTAACTTGAACCTGTGCTAGTTATCTCTACCGAAGTGA 133742
QY 1853 TGTGTAGTTTCTGTTTAAATTTCAAGCAACTGGAATAATCCATCTAATTTGCTT 1912
Db 133743 TGTGTAGTTTCTGTTTAAATTTCAAGCAACTGGAATAATCCATCTAATTTGCTT 133802
QY 1913 TCTTTCCCAAGAGTTTAAATGATATGATGCTGCTTCTAAATTTGAGACAAAGCTTAA 1972
Db 133803 TCTTTCCCAAGAGTTTAAATGATATGATGCTGCTTCTAAATTTGAGACAAAGCTTAA 133862
QY 1973 ATTGCAATGCAATCAATATATATTTTGTATAGTTTACAGTATACGAGTTGAGTATCC 2032
Db 133863 ATTGCAATGCAATCAATATATATTTTGTATAGTTTACAGTATACGAGTTGAGTATCC 133922
QY 2033 CTTTAGATGAGTCTGGACCAAGAGTCTTTTGGATTTTCAATTTTATTTTGGATTTTG 2092

QY	2242	TGCGC	CATTTGG	ACCACATCAGAAAGCAGAAAGTGTC	ACTATTTCAAGTCTAGTGTCTCAAAAAG	2301
DB	2655	TGCGC	CATTTGG	ACCACATCAGAAAGCAGAAAGTGTC	ACTATTTCAAGTCTAGTGTCTCAAAAAG	2714
QY	2302	TTTCAG	ATGTTAAG	CTTGGTGATGTCAGTTCATGCCAGTGATCCGAGTACTTTTGGGAAGCCA	2361	
DB	2715	TTTCAG	ATGTTAAG	CTTGGTGATGTCAGTTCATGCCAGTGATCCGAGTACTTTTGGGAAGCCA	2774	
QY	2362	AGCAG	TGTGATCTCT	TGAGCCACGAGTTTGAGGCCAGACTGCAACACACAGTGGAGACC	2421	
DB	2775	AGCAG	TGTGATCTCT	TGAGCCACGAGTTTGAGGCCAGACTGCAACACACAGTGGAGACC	2834	
QY	2422	TCGTTT	CTCAAA	TAAATTA	AAAAAATTAGCCAGGTGTGGTGCAACCTGTAGTCCGAG	2481
DB	2835	TCGTTT	CTCAAA	TAAATTA	AAAAAATTAGCCAGGTGTGGTGCAACCTGTAGTCCGAG	2894
QY	2482	GTACT	CAGGAGGCTGAGGTAGTAGTAGTA	TGTTTGAGACTGGGAGGTGAGGCTGAACTGAG	2541	
DB	2895	GTACT	CAGGAGGCTGAGGTAGTAGTAGTA	TGTTTGAGACTGGGAGGTGAGGCTGAACTGAG	2955	
QY	2542	CCAGG	ATCTTGCC	ACCACATTTCCAGCTTGGGCAACAGAGTGAGACCCCTGTCTC	2596	
DB	2955	CCAGG	ATCTTGCC	ACCACATTTCCAGCTTGGGCAACAGAGTGAGACCCCTGTCTC	3014	
QY	2597	AAAAA	AAAAAAGTTTCAGATTTTGGAGCA	TTTTTCAGATCTTCAGATAGGGATTTTCAAC	2656	
DB	3015	AAAAA	AAAAAAGTTTCAGATTTTGGAGCA	TTTTTCAGATCTTCAGATAGGGATTTTCAAC	3074	
QY	2657	CTGTACT	GACCTTTTAGTCA	TTTGACAAGCATTAATCAATAGGTGGACTCCAGATACTCA	2716	
DB	3075	CTGTACT	GACCTTTTAGTCA	TTTGACAAGCATTAATCAATAGGTGGACTCCAGATACTCA	3134	
QY	2717	TTTGCT	GTATACAT	TTTGCCCTCTCTATTCAACGAATCTTATATGCCCTCTTGTGGTGAT	2776	
DB	3135	TTTGCT	GTATACAT	TTTGCCCTCTCTATTCAACGAATCTTATATGCCCTCTTGTGGTGAT	3194	
QY	2777	TTTTAAT	GTGCGGAAGGGAACAAATAGAAAA	TTTTTGCAATCTAGAAAAAGTCATTTCTGTCAA	2836	
DB	3195	TTTTAAT	GTGCGGAAGGGAACAAATAGAAAA	TTTTTGCAATCTAGAAAAAGTCATTTCTGTCAA	3254	
QY	2837	AAATAT	GTCAGTCCCTGTAGATATATAGCCAA	TTTTTAGGAAATGACAAAAATTTTTCATCTTT	2896	
DB	3255	AAATAT	GTCAGTCCCTGTAGATATATAGCCAA	TTTTTAGGAAATGACAAAAATTTTTCATCTTT	3314	
QY	2897	CGTCTG	CCCTTGTAGCTGTTTATATGATA	TAAATACCTTATTTGTATATAAAATTAATTTTA	2956	
DB	3315	CGTCTG	CCCTTGTAGCTGTTTATATGATA	TAAATACCTTATTTGTATATAAAATTAATTTTA	3374	
QY	2957	ATTTGAG	TAACAATCTGGAATTA	TATCAGAGAAGGGGCAAGCAATAGGTTAATAAACAGTAT	3016	
DB	3375	ATTTGAG	TAACAATCTGGAATTA	TATCAGAGAAGGGGCAAGCAATAGGTTAATAAACAGTAT	3434	
QY	3017	TGATTTG	TAGAGGAACGTTGAAATCCAGAGCATCA	TGTCTCTGTGGTGTTTCAACATA	3076	
DB	3435	TGATTTG	TAGAGGAACGTTGAAATCCAGAGCATCA	TGTCTCTGTGGTGTTTCAACATA	3494	
QY	3077	AGCCAC	ACAGATGCTCTTAATCTTTTCCGAGATCT	TAGTTTTTCAGCAAGCAGGATTTAAG	3136	
DB	3495	AGCCAC	ACAGATGCTCTTAATCTTTTCCGAGATCT	TAGTTTTTCAGCAAGCAGGATTTAAG	3554	
QY	3137	AAATG	TAACTACTCTTATGTGTTATGAGAA	CAATAGAACTATTTGCTGTATAAGTGCTTT	3196	
DB	3555	AAATG	TAACTACTCTTATGTGTTATGAGAA	CAATAGAACTATTTGCTGTATAAGTGCTTT	3614	
QY	3197	TTAACTG	CTGTAAATTTTGTGAAGCTTA	CTTTTATGCAATATAAATTTTGAACATTTTACA	3256	
DB	3615	TTAACTG	CTGTAAATTTTGTGAAGCTTA	CTTTTATGCAATATAAATTTTGAACATTTTACA	3674	
QY	3257	TTGTTT	ATATTTTTTAATCAGTTTTTACTCAAGTGTGA	TTATATACAAGAAATGTAAACAC	3316	
DB	3675	TTGTTT	ATATTTTTTAATCAGTTTTTACTCAAGTGTGA	TTATATACAAGAAATGTAAACAC	3734	
QY	3317	TGTAAG	GGGTAGAGTTATTAAGAA	TTTTTGTCAAAATGTATTCACCCATGTAGTCAACCTCTTA	3376	

Db	3735	TGTAAGGTTAGATTATAGAATTTTGTCAATGTATTCACCCATGTAGTCACCTCCTTA	3794
Qy	3377	TGAAGACACAGAACACGTTACATCCTCCCGAAGTTCCACAGTGCTCCTTTCCCTGAGT	3436
Db	3795	TGAAGACACAGAACACGTTACATCCTCCCGAAGTTCCACAGTGCTCCTTTCCCTGAGT	3854
Qy	3437	TTCCACAGTCTCGGCAACCAATGATCGTTCCTGTATAATTTATAACCTGTTCTAGATATTG	3496
Db	3855	TTCCACAGTCTCGGCAACCAATGATCGTTCCTGTATAATTTATAACCTGTTCTAGATATTG	3914
Qy	3497	TAGCAATGTACCCCTTTCCATATTTATTTTGTGTGTGAAGCTTCTTTTAGTCATTATAA	3556
Db	3915	TAGCAATGTACCCCTTTCCATATTTATTTTGTGTGTGAAGCTTCTTTTAGTCATTATAA	3974
Qy	3557	TATTTTGGAGATTCATCTATGTGTTTAATGTTCTATCAGTAGTTGTGTACATCTTACTTGCTC	3616
Db	3975	TATTTTGGAGATTCATCTATGTGTTTAATGTTCTATCAGTAGTTGTGTACATCTTACTTGCTC	4034
Qy	3617	AGCATATCCACATATAGATATACTATAATTTGTTTAATCTAATCACCTGATGGATATGTAGG	3676
Db	4035	AGCATATCCACATATAGATATACTATAATTTGTTTAATCTAATCACCTGATGGATATGTAGG	4094
Qy	3677	ATATTTAAGTTTTTGACATTTATGAATAAAGTGCGCTATAAATG	3718
Db	4095	ATATTTAAGTTTTTGACATTTATGAATAAAGTGCGCTATAAATG	4136

RESULT 5	AL606804	185467 bp	DNA	linear	PRI 23-OCT-2002
AL606804					
LOCUS					
DEFINITION	Human DNA sequence from clone RP11-978115 on chromosome 1, complete sequence.				
ACCESSION	AL606804				
VERSION	AL606804.11	GI:24366459			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 185467)				
AUTHORS	Almeida,J.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Oct 24, 2002 this sequence version replaced gi:18121511.				
COMMENT					

COMMENT

On Oct 24, 2002 this sequence version replaced gi:18121511.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one W13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em1, EMBL; Sw1,
SWISSPROT; Tr1, TREMBL; Wp1, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chrl>
RP11-978I15 is from the library RP11-11.4 constructed by the group
of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBACe3.6

Db 2835 TCGTTTCTACAAATTAATAAATATACCAAGTGTGGTGCACACTGTAGTCCAG 2894
 Qy 2482 GTACTCAGAGGCTGAGGTAGTAGGATTTGAGACTGGAGGTTGAGGCTGAACTGAG 2541
 Db 2895 GTACTCAGAGGCTGAGGTAGTAGGATTTGAGACTGGAGGTTGAGGCTGAACTGAG 2954
 Qy 2542 CCAGATCTTGCACCAATTCAGCTTGGGCAACAGAGTGAGACCCCTGTCTC-----AA 2596
 Db 2955 CCAGATCTTGCACCAATTCAGCTTGGGCAACAGAGTGAGACCCCTGTCTC-----AA 3014
 Qy 2597 AAAAAAAGTTTTCAGATTTTGGAGCAATTCAGATCTTTCAGATTTAGGCAATTTTCAAC 2656
 Db 3015 AAAAAAAGTTTTCAGATTTTGGAGCAATTCAGATCTTTCAGATTTAGGCAATTTTCAAC 3074
 Qy 2657 CTGTACTGACCTTTTGTAGTCAATGCAAGCAATTAATCAATGAGTGGACTCCAGATCACTCA 2716
 Db 3075 CTGTACTGACCTTTTGTAGTCAATGCAAGCAATTAATCAATGAGTGGACTCCAGATCACTCA 3134
 Qy 2717 TTGTCTGTATACAGATTTTGTCTCTCTATTCAGCAATTCCTATGCTTGTGGTGTAT 2776
 Db 3135 TTGTCTGTATACAGATTTTGTCTCTCTATTCAGCAATTCCTATGCTTGTGGTGTAT 3194
 Qy 2777 TTTAATGTCGGAAGGAAACAAATAGAAATTTTGAATTTCTAGAAAAGTCATTTCTCTCAA 2836
 Db 3195 TTTAATGTCGGAAGGAAACAAATAGAAATTTTGAATTTCTAGAAAAGTCATTTCTCTCAA 3254
 Qy 2837 AATATGTCAGTCTGTAGATATTAGCCAAATTTAGGAAATGACAAAATTTTCTCTTTT 2896
 Db 3255 AATATGTCAGTCTGTAGATATTAGCCAAATTTAGGAAATGACAAAATTTTCTCTTTT 3314
 Qy 2897 CGTCTGCTTTGTAGTCTTTTATGATATAATACCTTATTTGTAATAAAATTAATTTTA 2956
 Db 3315 CGTCTGCTTTGTAGTCTTTTATGATATAATACCTTATTTGTAATAAAATTAATTTTA 3374
 Qy 2957 ATTGAGTAAACATCTGGAATTTATCAGAAAGGCGCAAGCAATAGTTTAATAAAGATAT 3016
 Db 3375 ATTGAGTAAACATCTGGAATTTATCAGAAAGGCGCAAGCAATAGTTTAATAAAGATAT 3434
 Qy 3017 TGATTTGTTAGAGGAAAGTGTGAAATCCAAAGAGCATCAATGTCTTCGTGGTTTACCAATA 3076
 Db 3435 TGATTTGTTAGAGGAAAGTGTGAAATCCAAAGAGCATCAATGTCTTCGTGGTTTACCAATA 3494
 Qy 3077 AGCCACAGAGATGCTTAAATCTTTCCGAGATCTAGTTTTCAGCAAGACGAGATTTAG 3136
 Db 3495 AGCCACAGAGATGCTTAAATCTTTCCGAGATCTAGTTTTCAGCAAGACGAGATTTAG 3554
 Qy 3137 AAATGTAACATCTTATGTTGTTATGAAGAAACAAATAGAAATCAATGCTGTATAGTCTTT 3196
 Db 3555 AAATGTAACATCTTATGTTGTTATGAAGAAACAAATAGAAATCAATGCTGTATAGTCTTT 3614
 Qy 3197 TTAACCTGTAATTTTGTGAAGCTTATCTTTTATGCAATTAATAATTTGAAATTTTACA 3256
 Db 3615 TTAACCTGTAATTTTGTGAAGCTTATCTTTTATGCAATTAATAATTTGAAATTTTACA 3674
 Qy 3257 TTGTTTATATTTTAACTAGTTTACTCAAGTGTGATTTATACAGAAAATGTAACCAAC 3316
 Db 3675 TTGTTTATATTTTAACTAGTTTACTCAAGTGTGATTTATACAGAAAATGTAACCAAC 3734
 Qy 3317 TGTAGGGTAGAGTTTATAGAAATTTTGTCAAATGTTATCACCATGTTAGTCACTCTTTA 3376
 Db 3735 TGTAGGGTAGAGTTTATAGAAATTTTGTCAAATGTTATCACCATGTTAGTCACTCTTTA 3794
 Qy 3377 TGAAGAGACAGACGATCATCTCCAGAAAGTTTCCAGAGTGTCTTTTCCCTGAGT 3436
 Db 3795 TGAAGAGACAGACGATCATCTCCAGAAAGTTTCCAGAGTGTCTTTTCCCTGAGT 3854
 Qy 3437 TTCACAGTCTCGGCAACCAATCATCTGCTGCTATATAATTTATTAATCTGTTCTAGATTTTG 3496
 Db 3855 TTCACAGTCTCGGCAACCAATCATCTGCTGCTATATAATTTATTAATCTGTTCTAGATTTTG 3914
 Qy 3497 TAGCAATGACCTTTTCCATATTTATTTTGTGTGTGTAAGGCTTCTTTTAGTCAATTATAA 3556

Db 3915 TAGCAATGACCTTTCCATATTTATTTTGTGTGTGTAAGGCTTCTTTTAGTCAATTATAA 3974
 Qy 3557 TATTTTTCAGATTTCACTCTATGTTTAAATGTTTCTATCAGTAGTGTGATCACTTACTTGTCTC 3616
 Db 3975 TATTTTTCAGATTTCACTCTATGTTTAAATGTTTCTATCAGTAGTGTGATCACTTACTTGTCTC 4034
 Qy 3617 AGCATATCACCATATAGATATACATAATTTGTTTAAATCTTAATCTGATGGATATGTAGG 3676
 Db 4035 AGCATATCACCATATAGATATACATAATTTGTTTAAATCTTAATCTGATGGATATGTAGG 4094
 Qy 3677 ATATTAAAGTTTTCAGATTTGATTAATAAGTGGCTATATAAG 3718
 Db 4095 ATATTAAAGTTTTCAGATTTGATTAATAAGTGGCTATATAAG 4136

RESULT 4
 AK124520 4136 bp mRNA linear PRI 09-SEP-2003
 LOCUS Homo sapiens cDNA FLJ42529 fis, clone BRACE3002390.
 DEFINITION AK124520
 ACCESSION AK124520
 VERSION AK124520.1 GI:34530324
 KEYWORDS oligo capping, fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., Kamiyama, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project

TITLE Unpublished
 JOURNAL 2 (bases 1 to 4136)
 REFERENCE Isogai, T. and Yamamoto, J.
 AUTHORS Direct Submission
 TITLE Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 JOURNAL NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
 source Location/Qualifiers
 1. 4136
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRACE3002390"
 /tissue_type="cerebellum"
 /clone_lib="BRACE3"
 /note="cloning vector: pME18SFL3"
 136..537
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAC85873.1"
 /db_xref="GI:34530325"
 /translation="MVMLEDSSSTCGGARNVCELVFLSVLGCQSKRGQMTQQAQ RMLRAGREASSETNPEGLTILDFQSEPLAARGWQEHFVRCNVLPHPAFSWCFQG NFPKWSQELNRNVTKRWELIISSENSGRK"
 CDS

Query Match 95.4%; Score 3590; DB 9; Length 4136;
 Best Local Similarity 99.3%; Pred. No. 0;

QY 262 TCCTCCCTGAGTCCCAATGATGGCTATGAGAACATTTGACTCCCTCACAAGGAAAGTG 321
DB 675 TCCTCCCTGAGTCCCAATGATGGCTATGAGAACATTTGACTCCCTCACAAGGAAAGTG 734
QY 322 AGACAGTTTATAGAGAAAGTTCAGAGACAGAAATATGCCCTTTCTTAGGACTTCTGTTAGTAGG 381
DB 735 AGACAGTTTATAGAGAAAGTTCAGAGACAGAAATATGCCCTTTCTTAGGACTTCTGTTAGTAGG 794
QY 382 CCTTGTCTCTGCA CCCATGAGCATGATATGAAGTTGTGTTTCCACACTAAATCCTCAA 441
DB 795 CCTTGTCTCTGCA CCCATGAGCATGATATGAAGTTGTGTTTCCACACTAAATCCTCAA 854
QY 442 GCTGCTTTATCACCTTCCAGCAATGAGACAAATGAGACATGAGACATGAGACATGAGGCAAGTTG 501
DB 855 GCTGCTTTATCACCTTCCAGCAATGAGACAAATGAGACATGAGACATGAGGCAAGTTG 914
QY 502 TTCAACCTGAGCAGTGCATGAAAATTCCTTCTGCTGCTAAAGTTTAGAAATATTATCTTA 561
DB 915 TTCAACCTGAGCAGTGCATGAAAATTCCTTCTGCTGCTAAAGTTTAGAAATATTATCTTA 974
QY 562 TTATATATCTTAGGCAACTCTGATATGTGGCATCTCTGTGGCTTAGTGAATCATAGA 621
DB 975 TTATATATCTTAGGCAACTCTGATATGTGGCATCTCTGTGGCTTAGTGAATCATAGA 1034
QY 622 AATTGACACATGACCTAAATAATTTCTTAGGCAATAATGTAATAATGAAATTTCTAATCCC 681
DB 1035 AATTGACACATGACCTAAATAATTTCTTAGGCAATAATGTAATAATGAAATTTCTAATCCC 1094
QY 682 AGGTGATAAAAAAAATTTCTTAGGCAATAATGTAATAATGAAATTTCTAATCCC 741
DB 1095 AGGTGATAAAAAAAATTTCTTAGGCAATAATGTAATAATGAAATTTCTAATCCC 1154
QY 742 CCTGACTAACTGATGACCTCTTCTAGGCCAAAGAGACCTCAGATGAACCTGAAAGAC 801
DB 1155 CCTGACTAACTGATGACCTCTTCTAGGCCAAAGAGACCTCAGATGAACCTGAAAGAC 1214
QY 802 TGAATTCCTGGCCATGATAGGAAGGAGGTGAGACACACCTTGTATACCCCTTCCCTTTT 861
DB 1215 TGAATTCCTGGCCATGATAGGAAGGAGGTGAGACACACCTTGTATACCCCTTCCCTTTT 1274
QY 862 GAGTTTATGCAAGTGACAGGATGAGTATGAAGACTGATGAATPAGACTGATTTGG 921
DB 1275 GAGTTTATGCAAGTGACAGGATGAGTATGAAGACTGATGAATPAGACTGATTTGG 1334
QY 922 CAATAAGAGTCCCAATTCGAACCTGACTCTGGTGTAGATCACACCTGCTGAGGATTC 981
DB 1335 CAATAAGAGTCCCAATTCGAACCTGACTCTGGTGTAGATCACACCTGCTGAGGATTC 1394
QY 982 CATCTATGAGACTTTGCTACATAACAGAGACCTTGGTTTCCACACCCCTTTATTTTAG 1041
DB 1395 CATCTATGAGACTTTGCTACATAACAGAGACCTTGGTTTCCACACCCCTTTATTTTAG 1454
QY 1042 CTAAGCAATCTTTCTGCTGCTTAAAGTCTTTAGACAAAGCTTAACTCTTTCAACC 1101
DB 1455 CTAAGCAATCTTTCTGCTGCTTAAAGTCTTTAGACAAAGCTTAACTCTTTCAACC 1514
QY 1102 AATTGCCAATCAGACAACTTTGAACTACCTATGACCTGTAGCTCTCTCTGCTTCAA 1161
DB 1515 AATTGCCAATCAGACAACTTTGAACTACCTATGACCTGTAGCTCTCTCTGCTTCAA 1574
QY 1162 GATCTTGGCTCTTTAAAGTGAACCGATGTGCACTTTTCCATTTTATGATTTATGCTCTTGG 1221
DB 1575 GATCTTGGCTCTTTAAAGTGAACCGATGTGCACTTTTCCATTTTATGATTTATGCTCTTGG 1634
QY 1222 TTGTAACCTCTGCTCTCCCTAAATGTATAAAGTAAACGGTGAACCTGACCACTCAGGCA 1281
DB 1635 TTGTAACCTCTGCTCTCCCTAAATGTATAAAGTAAACGGTGAACCTGACCACTCAGGCA 1694
QY 1282 CACTTTCTCAGGACCTCTGAGAGTGTATCCAGGCCATGGTAAGTCATGTTGGCTCAGA 1341
DB 1695 CACTTTCTCAGGACCTCTGAGAGTGTATCCAGGCCATGGTAAGTCATGTTGGCTCAGA 1754

QY 1342 ATCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTTTGGTTTACCAATTAAGTCTCCAAA 1401
DB 1755 ATCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTTTGGTTTACCAATTAAGTCTCCAAA 1814
QY 1402 TATATGTCGAAGATCTTCAATTCGAAGCTGCTCACCAGAAATTTCAATGCGCAATCTC 1461
DB 1915 TATATGTCGAAGATCTTCAATTCGAAGCTGCTCACCAGAAATTTCAATGCGCAATCTC 1874
QY 1462 CCCATCAATTTACCTATTTTCATCTTTGAGGTGTAATCTACTCAATAAATCTGTGTAAGACC 1521
DB 1875 CCCATCAATTTACCTATTTTCATCTTTGAGGTGTAATCTACTCAATAAATCTGTGTAAGACC 1934
QY 1522 AGTGACAGACCCCTTTGCTAAACCTGACATTTACTCAATTTTCTTTTCTATGTACTGG 1581
DB 1935 AGTGACAGACCCCTTTGCTAAACCTGACATTTACTCAATTTTCTTTTCTATGTACTGG 1994
QY 1582 ATATTTTGTGATATAAATTTGAGTGTGAGTGTGTAATCTCAAAATTTAATAGTTTTCATGTTGGCT 1641
DB 1995 ATATTTTGTGATATAAATTTGAGTGTGAGTGTGTAATCTCAAAATTTAATAGTTTTCATGTTGGCT 2054
QY 1642 TTTCTGAGAGAGAAATTTGAAAGTGTCAAAAAATAAAAAAGATGAAATGAAGCATATAT 1701
DB 2055 TTTCTGAGAGAGAAATTTGAAAGTGTCAAAAAATAAAAAAGATGAAATGAAGCATATAT 2114
QY 1702 AATTGTCAATTTTTCATTTTCTAGTCAACAGAGAAATCGAAGGATTCGTTTCAATTAAT 1761
DB 2115 AATTGTCAATTTTTCATTTTCTAGTCAACAGAGAAATCGAAGGATTCGTTTCAATTAAT 2174
QY 1762 AGTAAAAATTTGAAATTAACCTGCTGCTATATTTGTTTGCACACACTAGTAAATTTAA 1821
DB 2175 AGTAAAAATTTGAAATTAACCTGCTGCTATATTTGTTTGCACACACTAGTAAATTTAA 2234
QY 1822 CTTGTGACTAGTTTATCTCTACGAAGGTGATGTAGTTTCTGTTTTTAAAAATTCAGAC 1881
DB 2235 CTTGTGACTAGTTTATCTCTACGAAGGTGATGTAGTTTCTGTTTTTAAAAATTCAGAC 2294
QY 1882 AAATCGGAAAAATTAATCCATCTAATTAATGCTTTCTTCCCAAGGTTTTTAAATGATATG 1941
DB 2295 AAATCGGAAAAATTAATCCATCTAATTAATGCTTTCTTCCCAAGGTTTTTAAATGATATG 2354
QY 1942 CAGCTTCTCTAAATTTGAGACAAAAAGCTTTAATTTGACAAATGCAATTCATTTATATTTT 2001
DB 2355 CAGCTTCTCTAAATTTGAGACAAAAAGCTTTAATTTGACAAATGCAATTCATTTATATTTT 2414
QY 2002 TGTATAGTTACAGTATACGAGTTGAGTATCCCTTAGATGAGATGCTTGGACACAGAGTG 2061
DB 2415 TGTATAGTTACAGTATACGAGTTGAGTATCCCTTAGATGAGATGCTTGGACACAGAGTG 2474
QY 2062 TTTTGGATTTTCAGATTTTATTTTGGATTTTGGATTTTGGATTTTCCATATATTAATGAGAGT 2121
DB 2475 TTTTGGATTTTCAGATTTTATTTTGGATTTTGGATTTTGGATTTTCCATATATTAATGAGAGT 2534
QY 2122 TGGAAAAATGGGATTTCAAGTCTAATCAATAAATTCACCTTAATTTGTTGATATACCTTTATCT 2181
DB 2535 TGGAAAAATGGGATTTCAAGTCTAATCAATAAATTCACCTTAATTTGATATACCTTTATCT 2594
QY 2182 GAATAGCTTGAAGGTAATTTTATACAAATTTTAAATAAATTTTATGCTTGAACAGAGTT 2241
DB 2595 GAATAGCTTGAAGGTAATTTTATACAAATTTTAAATAAATTTTATGCTTGAACAGAGTT 2654
QY 2242 TCGGCACATTTGGACCATCAGAAAGCAGAGTGTCTACTATTTCAAGTCAGTGTCTAAAAAG 2301
DB 2655 TCGGCACATTTGGACCATCAGAAAGCAGAGTGTCTACTATTTCAAGTCAGTGTCTAAAAAG 2714
QY 2302 TTTTCAGATTTTAAAGCTGTTGATGCTAGTTTCTATCCAGTGTATCCGAGTACTTTTGGGAGCCA 2361
DB 2715 TTTTCAGATTTTAAAGCTGTTGATGCTAGTTTCTATCCAGTGTATCCGAGTACTTTTGGGAGCCA 2774
QY 2362 AGACAGGTGGATCTCTTAGGCCAGAGTTTGAAGCCAGACTSCACAAACACAGTGAAGCC 2421
DB 2775 AGACAGGTGGATCTCTTAGGCCAGAGTTTGAAGCCAGACTSCACAAACACAGTGAAGCC 2834
QY 2422 TCGTTTCTACAAATAATTAATAAAAAATTAGCCAGGTGTGGTGGACACCTGTAGTCCCG 2481

Db	2414	TGGAATAAGGGATTCAAGTCTAAATCATATAAAATTTCACTTAATGTTGATATACACCTTATCT	2473
Qy	2182	GAATAGCCTGAAGGTAAATTTTATACAAATATTTTAAATAATATTTATGCTCTGAACACAGATT	2241
Db	2474	GAATAGCCTGAAGGTAAATTTTATACAAATATTTTAAATAATATTTATGCTCTGAACACAGATT	2533
Qy	2242	TGGGCACATTTGGACCAATCAGAAACGAGAAGTGTCATTTTCAAGTCAGTGCTCAAAAAG	2301
Db	2534	TGGGCACATTTGGACCAATCAGAAACGAGAAGTGTCATTTTCAAGTCAGTGCTCAAAAAG	2593
Qy	2302	TTTCAGATGTTAAGCTGGTGATCGAGTTCATGCCAGTGATCCGAGTACTTTGGGAGGCCA	2361
Db	2594	TTTCAGATGTTAAGCTGGTGATCGAGTTCATGCCAGTGATCCGAGTACTTTGGGAGGCCA	2653
Qy	2362	AGACAGGTGGATCTCTTTGAGCCACGAGTTTGAAGCCAGACTGCACAAACAGTGAAGACC	2421
Db	2654	AGACAGGTGGATCTCTTTGAGCCACGAGTTTGAAGCCAGACTGCACAAACAGTGAAGACC	2713
Qy	2422	TCGTTTCTCAAAATAATTAATAAAATTAGCCAGGTGTGGTGGTGCACTGTAGTCCCAAG	2488
Db	2714	TCGTTTCTCAAAATAATTAATAAAATTAGCCAGGTGTGGTGGTGCACTGTAGTCCCAAG	2773
Qy	2482	GTACTCAGGAGGCTGAGGTAGTAGGATTTGTTTGAGACTGGGAGGTTGAGCCTGMACTGAG	2541
Db	2774	GTACTCAGGAGGCTGAGGTAGTAGGATTTGTTTGAGACTGGGAGGTTGAGCCTGMACTGAG	2833
Qy	2542	CCAGGATCTTGCCACACACATTCACAGCTTTGGGCAACAGAGTGAGACCTGCTGTC - AAAAAA	2600
Db	2834	CCAGGATCTTGCCACACACATTCACAGCTTTGGGCAACAGAGTGAGACCTGCTGTC - AAAAAA	2893
Qy	2601	AAAAAAAGTTTCAGATTTTTGGAGCATTTTCAGATCTTTCAGATTAGGATTTTCAACCTGT	2660
Db	2894	AAAAAAAGTTTCAGATTTTTGGAGCATTTTCAGATCTTTCAGATTAGGATTTTCAACCTGT	2953
Qy	2661	ACTGACCTTTTGTGTCATGTGACAGCATTAATCAATAGGTGGACTCCAGATTAAGTCAATTTG	2720
Db	2954	ACTGACCTTTTGTGTCATGTGACAGCATTAATCAATAGGTGGACTCCAGATTAAGTCAATTTG	3013
Qy	2721	CTGTATACACATTTTGCTCTCTATTCACACGAATTTCTATGCCCTCTTGTGGTGATTTTA	2780
Db	3014	CTGTATACACATTTTGCTCTCTATTCACACGAATTTCTATGCCCTCTTGTGGTGATTTTA	3073
Qy	2781	ATGTGCGGAGGGGAAAACAATAGAAATTTTGCAATTTCTAGAAAAAGTCATCTGTCAAAATA	2840
Db	3074	ATGTGCGGAGGGGAAAACAATAGAAATTTTGCAATTTCTAGAAAAAGTCATCTGTCAAAATA	3133
Qy	2841	TGTGAGTCTGTAGATATTAGCCAAATTTAGGAAAATTCACAAAATTTTTTTTACTTTTCGTC	2900
Db	3134	TGTGAGTCTGTAGATATTAGCCAAATTTAGGAAAATTTTCACAAAATTTTTTTTACTTTTCGTC	3193
Qy	2901	TGCCCTTTGTAGCTGTTTTATCATATAAATACCTTATTTGTAAATAAAATTAATTTTAATTT	2960
Db	3194	TGCCCTTTGTAGCTGTTTTATCATATAAATACCTTATTTGTAAATAAAATTAATTTTAATTT	3253
Qy	2961	GAGTAAACAATCTGGAATTTATCAGAGAGGGGCAAGCAATAGGTTTAATAACAGTATTGAT	3020
Db	3254	GAGTAAACAATCTGGAATTTATCAGAGAGGGGCAAGCAATAGGTTTAATAACAGTATTGAT	3313
Qy	3021	TGGTAGAAGGAAAGGTTGAAATCCAAAGAGCATCAATGTCTTCTGGTGGTTTCAACATAAGCC	3080
Db	3314	TGGTAGAAGGAAAGGTTGAAATCCAAAGAGCATCAATGTCTTCTGGTGGTTTCAACATAAGCC	3373
Qy	3081	ACAGCAGATGCTTAAATCTTTCCGAGATCTAGTTTTTTTTCAGCAAAAGCAGGATTTAAGAAAT	3140
Db	3374	ACAGCAGATGCTTAAATCTTTCCGAGATCTAGTTTTTTTTCAGCAAAAGCAGGATTTAAGAAAT	3433
Qy	3141	GTAACTATCTTATGTGGTTTATGAAGAAACAATAGAAATCATTCGTATAGTGCTTTTTTAA	3200
Db	3434	GTAACTATCTTATGTGGTTTATGAAGAAACAATAGAAATCATTCGTATAGTGCTTTTTTAA	3493
Qy	3201	CCTGTAAATTTTGTGAAGCTTATCTTTTATGCATATAAATAATTTTGAACATTTTACATTTGT	3260
Db	3494	CCTGTAAATTTTGTGAAGCTTATCTTTTATGCATATAAATAATTTTGAACATTTTACATTTGT	3553

QY	3261	TTATATTTTTAAATCAGTGTCTTCTCAAGTGTGATATATATACAGAAATGTAAACCACTGTA	3322
DB	3554	TTATATTTTTAAATCAGTGTCTTCTCAAGTGTGATATATATACAGAAATGTAAACCACTGTA	3613
QY	3321	AGGTGAGTGTATAAAGATTTTGTCAAAATGTATTACCCATGTAGTCACCTCTCTTATGAA	3380
DB	3614	AGGTGAGTGTATAAAGATTTTGTCAAAATGTATTACCCATGTAGTCACCTCTCTTATGAA	3673
QY	3381	GAGACAGAACACGTACATCTCTCCAGAAAGTTCCACAGTGTCTCTTTTCCCTGAGTTTCA	3440
DB	3674	GAGACAGAACACGTACATCTCTCCAGAAAGTTCCACAGTGTCTCTTTTCCCTGAGTTTCA	3733
QY	3441	CCAGTCTGGCAACCAATGATCTGCTTCGTATAATTAATAACTGTTCTAGATATTTGTGAGC	3500
DB	3734	CCAGTCTGGCAACCAATGATCTGCTTCGTATAATTAATAACTGTTCTAGATATTTGTGAGC	3793
QY	3501	AATGTACCCCTTCCATATTTATTTTGTGTGTGAAGCTTCTTTTAGTCAATATAATATT	3560
DB	3794	AATGTACCCCTTCCATATTTATTTTGTGTGTGAAGCTTCTTTTAGTCAATATAATATT	3853
QY	3561	TTTGAGATTCATCTATGTTTAAATGTTCTATCAGTAGTTGTACATCTTGTCTCAGCA	3620
DB	3854	TTTGAGATTCATCTATGTTTAAATGTTCTATCAGTAGTTGTACATCTTGTCTCAGCA	3913
QY	3621	TATCACCATATAGATATACTATAATTTGTTAACTTAATCACTGATGGATATGTAGGATAT	3680
DB	3914	TATCACCATATAGATATACTATAATTTGTTAACTTAATCACTGATGGATATGTAGGATAT	3973
QY	3681	TTAAGTTTTTGCACATTATGAATAAAGTGGCTATAAAATCAAAAAAATAAAAAA	3735
DB	3974	TTAAGTTTTTGCACATTATGAATAAAGTGGCTATAAAATCAAAAAAATAAAAAA	4028
RESULT 3			
LOCUS	Q0841482	4136 bp	DNA
DEFINITION	Sequence 129 from Patent EP1440981.	linear	PAT 02-AUG-2004
ACCESSION	Q0841482		
VERSION	Q0841482.1	GI:50893269	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S., Yamamoto,J., Isono,Y., Nagai,K. and Irie,R.		
TITLE	Full-length human cdna		
JOURNAL	Patent: EP 1440981-A 129 28-JUL-2004;		
FEATURES	Research Association for Biotechnology (JP)		
source	Location/Qualifiers		
	1..4136		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	95.4%;	Score 3590;	DB 6; Length 4136;
Best Local Similarity	99.3%;	Pred. No. 0;	
Matches 3617;	Conservative	0; Mismatches	20; Indels 5; Gaps 1
QY	82	AGAAAGGAAACCCAGATGAGAGAAAGAAACGGCAGGAAATGACTACATTTTGAAAGAAA	141
DB	495	AAAGATGGGAAATTTATCTCTCTCGAAACCTCAGCGAGGAAATGACTACATTTTGAAAGAAA	554
QY	142	CTTCAAGATCAAGATAAGAAAGCCAAAGATTTTCATCCACTTCTTAATCAGGAAACGAG	201
DB	555	CTTCAAGATCAAGATAAGAAAGCCAAAGATTTTCATCCACTTCTTAATCAGGAAACGAG	614
QY	202	AATGGCAGTGGTTCTGAAGAAGTGTGCTACACTGTCATTATCAATCCCATCAGAGA	261
DB	615	AATGGCAGTGGTTCTGAAGAAGTGTGCTACACTGTCATTATCAATCCCATCAGAGA	674

ORIGIN		/clone_lib="BRACE3"		/note="cloning vector: pME18SFL3"	
Query Match		95.6%; Score 3596.6; DB 9; Length 4032;			
Best Local Similarity		99.2%; Pred. No. 0;			
Matches 3625; Conservative		0; Mismatches 29; Indels 1; Gaps 1;			
QY	82	AGAGAGGAAACCCAGATAGGAAAGAAACCGGAGGAAATGACTACATTGGAAGAAA	141	DB	1334
DB	374	AAGATGGGAAATATCTCTCGGAAACCTCAGGAGGAAATGACTACATTGGAAGAAA	433	QY	1102
QY	142	CTTCAAGATCAAGATAGAAAGCCAAAGAGTTTCATCCACTCTCTAATCAGGAAACGAG	201	DB	1394
DB	434	CTTCAAGATCAAGATAGAAAGCCAAAGAGTTTCATCCACTCTCTAATCAGGAAACGAG	493	QY	1162
QY	202	AATGGCAGTGGTTCTGAAGAGGTGTGTACACTGTCTAATTAATCACATCCCCCATCAGAG	261	DB	1454
DB	494	AATGGCAGTGGTTCTGAAGAGGTGTGTACACTGTCTAATTAATCACATCCCCCATCAGAG	553	QY	1222
QY	262	TCCTCCCTGAGCTCCCAATGATGCTATGAGAACTTACCTCTCTAATCAGGAAACGAG	321	DB	1513
DB	554	TCCTCCCTGAGCTCCCAATGATGCTATGAGAACTTACCTCTCTAATCAGGAAACGAG	613	QY	1281
QY	322	AGACAGTTTACAGAAAGGTTCAGAGACAGAAATATGCCCTCTTAGGACTTCTGTAGTAGG	381	DB	1513
DB	614	AGACAGTTTACAGAAAGGTTCAGAGACAGAAATATGCCCTCTTAGGACTTCTGTAGTAGG	673	QY	1401
QY	382	CTTGTCTCTGACCCCATGAGCATGATTATGAAGTTGTGTTTCCACATAAATTCCTCA	441	DB	1693
DB	674	CTTGTCTCTGACCCCATGAGCATGATTATGAAGTTGTGTTTCCACATAAATTCCTCA	733	QY	1461
QY	442	GCTGCTTTATCACCTTCCAGCAATGAAGCAATGAGAGATAGCAGACTCTGGCGAAGTTG	501	DB	1753
DB	734	GCTGCTTTATCACCTTCCAGCAATGAAGCAATGAGAGATAGCAGACTCTGGCGAAGTTG	793	QY	1521
QY	502	TTCACTCTGAGCAGTGCATGAACATTCCTCTCTGCGCTAAGTTTGAAGATATATCTTA	561	DB	1813
DB	794	TTCACTCTGAGCAGTGCATGAACATTCCTCTCTGCGCTAAGTTTGAAGATATATCTTA	853	QY	1581
QY	562	TTATATATCTTACGCAACTCTGATATGTCGCTCTCTGCTGCTAGGTGAAATCATAGA	621	DB	1873
DB	854	TTATATATCTTACGCAACTCTGATATGTCGCTCTCTGCTGCTAGGTGAAATCATAGA	913	QY	1641
QY	622	AATTGACAAATCACTTAAATATCTATGTGTTTCTGTTGTTTAAAGTTTGAAGATATATCTTA	681	DB	1933
DB	914	AATTGACAAATCACTTAAATATCTATGTGTTTCTGTTGTTTAAAGTTTGAAGATATATCTTA	973	QY	1701
QY	682	AGGTGATAAAAAAACTTTCTTAGGCAATATATGTAATAATGAAATTAATTTCTAATCCC	741	DB	1993
DB	974	AGGTGATAAAAAAACTTTCTTAGGCAATATATGTAATAATGAAATTAATTTCTAATCCC	1033	QY	1761
QY	742	CTGACTAATGAAATGGAACCTCTTAGGCCAAAGAGACCTCAGATGAACCTGAAAGAC	801	DB	2053
DB	1034	CTGACTAATGAAATGGAACCTCTTAGGCCAAAGAGACCTCAGATGAACCTGAAAGAC	1093	QY	1821
QY	802	TGAATTTCTGCCATGATAGAGAGGAGTGCAGACACACTTGTATACCCCTTCCCTTTT	861	DB	2113
DB	1094	TGAATTTCTGCCATGATAGAGAGGAGTGCAGACACACTTGTATACCCCTTCCCTTTT	1153	QY	1881
QY	862	GGAGTTTATGCAAGATGACCCAGGATGAGTCAATAGACTGATGAAATAGACTGATTTGCG	921	DB	2173
DB	1154	GGAGTTTATGCAAGATGACCCAGGATGAGTCAATAGACTGATGAAATAGACTGATTTGCG	1213	QY	1941
QY	922	CAATAGAGTCCCAATTCACCTGACTCTGGTGTAGATCACACTGTCTGAGGGAATTC	981	DB	2233
DB	1214	CAATAGAGTCCCAATTCACCTGACTCTGGTGTAGATCACACTGTCTGAGGGAATTC	1273	QY	2001
QY	982	CATCTATGAGACTTCTCTACATAACAGAGACCTTGTGTTTCCACACCCCTTTATTTAG	1041	DB	2293
DB	1274	CATCTATGAGACTTCTCTACATAACAGAGACCTTGTGTTTCCACACCCCTTTATTTAG	1333	QY	2061
QY	1042	CTAAGCATTTCTTCTACTGACTTCTTAAGTCTTTTAGACAAAGCTTAACCTCTTTCAACC	1101	DB	2353
				QY	2121
				DB	2413
				QY	2181

```
QY 2281 TTCAAGTCAGTGTCAAAAAGTTTCAGATGTTAAAGCTGGTGATGCAGTTCAATGCCAGTGA 2340
Db 2281 TTCAAGTCAGTGTCAAAAAGTTTCAGATGTTAAAGCTGGTGATGCAGTTCAATGCCAGTGA 2340
QY 2341 TCCGAGTACTTTGGGAAGCCCAAGACAGGTGGATCTCTTGAGCCAGGAGTTTGAGGCCAG 2400
Db 2341 TCCGAGTACTTTGGGAAGCCCAAGACAGGTGGATCTCTTGAGCCAGGAGTTTGAGGCCAG 2400
QY 2401 ACTGCAACAACAGTGTAGTCCAGTGTCTCTGTTCTCAAAATAATTAATAAATAGCCAGTGTGGT 2460
Db 2401 ACTGCAACAACAGTGTAGTCCAGTGTCTCTGTTCTCAAAATAATTAATAAATAGCCAGTGTGGT 2460
QY 2461 GGTGCACACCTGTAGTCCAGTGTCTCTGTTCTCAAAATAATTAATAAATAGCCAGTGTGGT 2520
Db 2461 GGTGCACACCTGTAGTCCAGTGTCTCTGTTCTCAAAATAATTAATAAATAGCCAGTGTGGT 2520
QY 2521 GGAGGTTGAGGCTGAACCTGAGCCAGGATCTTGCCACCACTTCAGCAGCTTGGGCCAACAGAG 2580
Db 2521 GGAGGTTGAGGCTGAACCTGAGCCAGGATCTTGCCACCACTTCAGCAGCTTGGGCCAACAGAG 2580
QY 2581 TGAGACCTGTCTCAAAAAGTTTCAGATTTTGAGGATTTTGAGGATTTTGAGGATTTTGAGGATTT 2640
Db 2581 TGAGACCTGTCTCAAAAAGTTTCAGATTTTGAGGATTTTGAGGATTTTGAGGATTTTGAGGATTT 2640
QY 2641 ATTAGGATTTTCAACCTGTACTGACCTTTTAGTCATTGACAGCAATTAATCAATAGGTG 2700
Db 2641 ATTAGGATTTTCAACCTGTACTGACCTTTTAGTCATTGACAGCAATTAATCAATAGGTG 2700
QY 2701 GACTCCAGATAACTCAATTTGCTGTATACACATTTTGCCTCTCTATTCAACGAATTTCTTAT 2760
Db 2701 GACTCCAGATAACTCAATTTGCTGTATACACATTTTGCCTCTCTATTCAACGAATTTCTTAT 2760
QY 2761 GCCCTCTTGTGGTAAATTTAATGTCGGAAGGAAACAAATAGAAATTTTGCAATTTCTAGA 2820
Db 2761 GCCCTCTTGTGGTAAATTTAATGTCGGAAGGAAACAAATAGAAATTTTGCAATTTCTAGA 2820
QY 2821 AAAGTCATTCTGCAAAATATGTCAGTCTGAGATATTAGCCAAATTTTAGGAAATGAC 2880
Db 2821 AAAGTCATTCTGCAAAATATGTCAGTCTGAGATATTAGCCAAATTTTAGGAAATGAC 2880
QY 2881 AAAATTTTACTTTTCGCTCTGCTTTAGTGTGTTTATGATAPAAATACCTTATTTGT 2940
Db 2881 AAAATTTTACTTTTCGCTCTGCTTTAGTGTGTTTATGATAPAAATACCTTATTTGT 2940
QY 2941 AATAAATAATTTTAAATTTGAGTAAACAAATCTGGAATATTACGAGAAGGGGCAACAAATA 3000
Db 2941 AATAAATAATTTTAAATTTGAGTAAACAAATCTGGAATATTACGAGAAGGGGCAACAAATA 3000
QY 3001 GGTAAATAACAGATATTGATTTGTTAGAGGAACGTTGAAATCCAGAGCATCAATGTCTT 3060
Db 3001 GGTAAATAACAGATATTGATTTGTTAGAGGAACGTTGAAATCCAGAGCATCAATGTCTT 3060
QY 3061 CTGGTGGTTTCAACATAAGCCACAGAGATGCTTAAATCTTTCCGAGATCTAGTTTTCAG 3120
Db 3061 CTGGTGGTTTCAACATAAGCCACAGAGATGCTTAAATCTTTCCGAGATCTAGTTTTCAG 3120
QY 3121 CAAGCAGATTTTAAAGTAAATGTAATCTTATGTTTATGAGGATTTTATGAGGATTTTATATAC 3180
Db 3121 CAAGCAGATTTTAAAGTAAATGTAATCTTATGTTTATGAGGATTTTATGAGGATTTTATATAC 3180
QY 3181 GCTGTATAAGTCTTTTAAACCTGTAATTTTGTGAAGCTTATCTTTTATGATATAAAT 3240
Db 3181 GCTGTATAAGTCTTTTAAACCTGTAATTTTGTGAAGCTTATCTTTTATGATATAAAT 3240
QY 3241 ATTGAACATTTTACATTTTATATTTTAAATCAAGTTTATCTCAAGTGTGATATATAC 3300
Db 3241 ATTGAACATTTTACATTTTATATTTTAAATCAAGTTTATCTCAAGTGTGATATATAC 3300
QY 3301 AAGAAATGTAACCTGTAAGGGTAGGTATACAAATTTTGTCAAAATGATTTTACCCCA 3360
Db 3301 AAGAAATGTAACCTGTAAGGGTAGGTATATAGAAATTTTGTCAAAATGATTTTACCCCA 3360
QY 3361 TGTAGTCACCTCTTATTAAGAGACAGACAGTATACCTCTCCAGAAAGTTTCCACAGTG 3420
```

```
Db 3361 TGTAGTCACCTCTTATTAAGAGACAGACAGTATACCTCTCCAGAAAGTTTCCACAGTG 3420
QY 3421 CTCCTTTTCCCTGAGTTTCCACAGTCTGGCAACCAATGATCTCTCGTATAATATAA 3480
Db 3421 CTCCTTTTCCCTGAGTTTCCACAGTCTGGCAACCAATGATCTCTCGTATAATATAA 3480
QY 3481 CTGTTCTAGATATTGTTAGCAATGATCCCTTTTCCATATTTATTTTGTGTGTAGGCTT 3540
Db 3481 CTGTTCTAGATATTGTTAGCAATGATCCCTTTTCCATATTTATTTTGTGTGTAGGCTT 3540
QY 3541 CTTTGTAGTCATTATAATATTGTTGAGATTCATCTATGTTTAAATGTTCTATCAGTAGTTGT 3600
Db 3541 CTTTGTAGTCATTATAATATTGTTGAGATTCATCTATGTTTAAATGTTCTATCAGTAGTTGT 3600
QY 3601 ACATCTTACTTGTCTCAGCATATCACCATATAGATATATCTTCTTAAATCTTAATCAATCA 3660
Db 3601 ACATCTTACTTGTCTCAGCATATCACCATATAGATATATCTTCTTAAATCTTAATCAATCA 3660
QY 3661 CTGATGGATATGTTAGGATATTAAAGTTTTCACATTTATGAATAAGTGGCTATAAATGAA 3720
Db 3661 CTGATGGATATGTTAGGATATTAAAGTTTTCACATTTATGAATAAGTGGCTATAAATGAA 3720
QY 3721 AAAAAAAGTATGTTAGGATATTAAAGTTTTCACATTTATGAATAAGTGGCTATAAATGAA 3762
Db 3721 AAAAAAAGTATGTTAGGATATTAAAGTTTTCACATTTATGAATAAGTGGCTATAAATGAA 3762
```

RESULT 2

AK126682
LOCUS Homo sapiens cDNA FLJ44728 fis, clone BRACE3024537. 4032 bp mRNA linear PRI 19-FEB-2004

DEFINITION

AK126682

ACCESSION

AK126682.1 GI:34533254

VERSION

oligo capping; fis (full insert sequence).

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,

Kodaira, H., Furiya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,

Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,

Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,

Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,

Kimura, K., Yanashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,

Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,

Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,

Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.

NEBO human cDNA sequencing project

Unpublished

2 (bases 1 to 4032)

Isogai, T. and Yamamoto, J.

Direct Submission

Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7

Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEBO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5' - & 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: HRI and

RAB; annotation: Reverse Proteomics Research Institute, HRI and

RAB.

Location/Qualifiers

1..4032

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRACE3024537"

/tissue_type="cerebellum"

FEATURES

source

Db 61 GGAGAGAAATCAAAAGAACCCAGAAAGGAAACCCAGATGAGGAAAGAAACGGCAGGAA 120
Qy 121 ATGACTACATTTGAAAGAAAATCTCAAGATCAAGATAGAAAGAACCCAGAGATTTTCATCC 180
Db 121 ATGACTACATTTGAAAGAAAATCTCAAGATCAAGATAGAAAGAACCCAGAGATTTTCATCC 180
Qy 181 ACTTCTAATCAGGAAACGAGAAATGGCAGTGGTCTCTGAAGAAAGTGTGCTACATGTCAAT 240
Db 181 ACTTCTAATCAGGAAACGAGAAATGGCAGTGGTCTCTGAAGAAAGTGTGCTACATGTCAAT 240
Qy 241 AATCACATCCCCATCAGAGATCTCCCTGAGCTCCCAATGATGATGCTATGAGAACAT 300
Db 241 AATCACATCCCCATCAGAGATCTCCCTGAGCTCCCAATGATGATGCTATGAGAACAT 300
Qy 301 GACTCCCTCAAGAGAAAGTGAGACAGTTTAGAGAAAGGTGAGAGACAGAAATATGCCCTT 360
Db 301 GACTCCCTCAAGAGAAAGTGAGACAGTTTAGAGAAAGGTGAGAGACAGAAATATGCCCTT 360
Qy 361 CTTAGGACTCTGTTAGTAGGCTTGTCTGACAGCTCTTATCACCTTCCAGCAATGAAGCAATGCAGAA 480
Db 361 CTTAGGACTCTGTTAGTAGGCTTGTCTGACAGCTCTTATCACCTTCCAGCAATGAAGCAATGCAGAA 480
Qy 481 TAGCAGACTCTGGGAAAGTGTTCACCTGAGCAGTGCATGAACATCTCTTCTGGCTA 540
Db 481 TAGCAGACTCTGGGAAAGTGTTCACCTGAGCAGTGCATGAACATCTCTTCTGGCTA 540
Qy 541 AAGTTTGAAGAAATATCTTATATATATATCTTATAGGCAATCTGATATGTTGGCATCTG 600
Db 541 AAGTTTGAAGAAATATCTTATATATATCTTATAGGCAATCTGATATGTTGGCATCTG 600
Qy 601 TGGCTTAGGTAATCATAGAAATTTGACAAATGACCTAAATATCTATGTTTTCG 660
Db 601 TGGCTTAGGTAATCATAGAAATTTGACAAATGACCTAAATATCTATGTTTTCG 660
Qy 661 TTGTAAGTTTGAAGACATGAGGTGATAAATAAATCTTCTTAGGCAATATGTAATA 720
Db 661 TTGTAAGTTTGAAGACATGAGGTGATAAATAAATCTTCTTAGGCAATATGTAATA 720
Qy 721 TGAATAATAATTTCTAATCCCTGACTAATCTGAATGACCTCTTCTAGGCCAAGAGA 780
Db 721 TGAATAATAATTTCTAATCCCTGACTAATCTGAATGACCTCTTCTAGGCCAAGAGA 780
Qy 781 CTTAGATGAACCTGAAAGACTGAAATCTGGCCATGATAGAAAGGAGGTGAGACACACC 840
Db 781 CTTAGATGAACCTGAAAGACTGAAATCTGGCCATGATAGAAAGGAGGTGAGACACACC 840
Qy 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGAACAGGATGAGTCAATAGACT 900
Db 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGAACAGGATGAGTCAATAGACT 900
Qy 901 GATGAATAGACTGATGTGGCAATAGAGTCCCAATTCACACCTGACTCTGTTGTAGAT 960
Db 901 GATGAATAGACTGATGTGGCAATAGAGTCCCAATTCACACCTGACTCTGTTGTAGAT 960
Qy 961 CACACACTGTCTGAGGATTCATCTATGAGACTTTGTCTACATAACAGAGACCTTGGTT 1020
Db 961 CACACACTGTCTGAGGATTCATCTATGAGACTTTGTCTACATAACAGAGACCTTGGTT 1020
Qy 1021 TCCACAACCCCTTATTTTAGCTTAAGCAATTCAGACAACTTTTGAATCTACCTATGACCT 1080
Db 1021 TCCACAACCCCTTATTTTAGCTTAAGCAATTCAGACAACTTTTGAATCTACCTATGACCT 1080
Qy 1081 CAAAGCTTAACTCTTTCAACCAATTCGCAATCAGACAACTTTGAATCTACCTATGACCT 1140
Db 1081 CAAAGCTTAACTCTTTCAACCAATTCGCAATCAGACAACTTTGAATCTACCTATGACCT 1140
Qy 1141 GTAGCTCTCTCTGCTTCAAGATCTTGCCTCTTTAAGCTGAACCGATGTGCACTTCCA 1200
Db 1141 GTAGCTCTCTCTGCTTCAAGATCTTGCCTCTTTAAGCTGAACCGATGTGCACTTCCA 1200

Qy 1201 TTTAATGATTTATGTCTTTGTTAACTCTCTCTCCCTAAATATGTAATAAGTAAACG 1260
Db 1201 TTTAATGATTTATGTCTTTGTTAACTCTCTCTCCCTAAATATGTAATAAGTAAACG 1260
Qy 1261 GTGACCTGACCACTCTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1320
Db 1261 GTGACCTGACCACTCTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1320
Qy 1321 GGTAAAGTCATGTTGGCTCAGAAATCAACCTCTTTAAATATTTTACAGAAATTTGGTTTGG 1380
Db 1321 GGTAAAGTCATGTTGGCTCAGAAATCAACCTCTTTAAATATTTTACAGAAATTTGGTTTGG 1380
Qy 1381 TTACCAATAGTCTCCCAAAATATATGTCTCCAAAGATCTTCAATTCCAAGCCTGCTCACCA 1440
Db 1381 TTACCAATAGTCTCCCAAAATATATGTCTCCAAAGATCTTCAATTCCAAGCCTGCTCACCA 1440
Qy 1441 AATTTCAAAATGCAACATCTCCCAATCTCCCAATCTTCAATTCCTTTGAGGTGTAACTTA 1500
Db 1441 AATTTCAAAATGCAACATCTCCCAATCTCCCAATCTTCAATTCCTTTGAGGTGTAACTTA 1500
Qy 1501 CTCAATAAACTGTGTAAAGCAGTGAACCCCTTTGCTTAACCTGACATTTTACTTCAAT 1560
Db 1501 CTCAATAAACTGTGTAAAGCAGTGAACCCCTTTGCTTAACCTGACATTTTACTTCAAT 1560
Qy 1561 TTTTCTTTTCTATGTAAGTATTTTTCATATAAACTTGCAGTAAATAGTTCACAAAAT 1620
Db 1561 TTTTCTTTTCTATGTAAGTATTTTTCATATAAACTTGCAGTAAATAGTTCACAAAAT 1620
Qy 1621 TAATAGTTTGTGACATTTGCTTTCTGAGAAGAAATGAAAGTGTCAAAAATAAAA 1680
Db 1621 TAATAGTTTGTGACATTTGCTTTCTGAGAAGAAATGAAAGTGTCAAAAATAAAA 1680
Qy 1681 AAGATGAAATGAAGCATATATATCTCAATTTTCTCAATTTTCTAGTCAACAGAGATC 1740
Db 1681 AAGATGAAATGAAGCATATATATCTCAATTTTCTCAATTTTCTAGTCAACAGAGATC 1740
Qy 1741 GAAGGATCTGTTTCAAAATATAGTAAAAATTTGAAAAATTAACCTTGTCTTATATTTGTTT 1800
Db 1741 GAAGGATCTGTTTCAAAATATAGTAAAAATTTGAAAAATTAACCTTGTCTTATATTTGTTT 1800
Qy 1801 GCAACACACTAGTAAATTTAACTGACTAGTATCTCTACCGAAGGTGGATGTAGT 1860
Db 1801 GCAACACACTAGTAAATTTAACTGACTAGTATCTCTACCGAAGGTGGATGTAGT 1860
Qy 1861 TTCTGCTTTTAAATTTCAAGCAAACTGAAAAATTAATCCATCTAATTTATGCTTTCTTCCC 1920
Db 1861 TTCTGCTTTTAAATTTCAAGCAAACTGAAAAATTAATCCATCTAATTTATGCTTTCTTCCC 1920
Qy 1921 AGAAGTTTAAATGATATGCGAGCTTCTTAATTTGGAGACAAAAAGCCTTAATGACAA 1980
Db 1921 AGAAGTTTAAATGATATGCGAGCTTCTTAATTTGGAGACAAAAAGCCTTAATGACAA 1980
Qy 1981 TGCAATTCATATATATTTTGTATAGTATAGATATAGTGTGAGTATCCCTTAGATG 2040
Db 1981 TGCAATTCATATATATTTTGTATAGTATAGATATAGTGTGAGTATCCCTTAGATG 2040
Qy 2041 AGATGCTTGGGACCGAGAGTGTGTTTGGATTTTTCAGATTTTATTTTGGATTTTGAATATTT 2100
Db 2041 AGATGCTTGGGACCGAGAGTGTGTTTGGATTTTTCAGATTTTATTTTGGATTTTGAATATTT 2100
Qy 2101 CCATACATATATAGAGAGTGTGGAATAATGGGATTCAGTCTAAATCATAAAAATTCATTTA 2160
Db 2101 CCATACATATATAGAGAGTGTGGAATAATGGGATTCAGTCTAAATCATAAAAATTCATTTA 2160
Qy 2161 TGTGTTGATATACCTTATCTGAATAGCTGAGGTAAATTTTATACAAATTTTAAATTA 2220
Db 2161 TGTGTTGATATACCTTATCTGAATAGCTGAGGTAAATTTTATACAAATTTTAAATTA 2220
Qy 2221 TTTTATGCTGAAAACAGAGTGTTCGCGACATTCGACCATCAGAAAGCAGAGTGTCACTAT 2280
Db 2221 TTTTATGCTGAAAACAGAGTGTTCGCGACATTCGACCATCAGAAAGCAGAGTGTCACTAT 2280

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2004, 12:52:13 : Search time 15375 Seconds
(without alignments)
11570.982 Million cell updates/sec

Title: US-10-005-907-1
Perfect score: 3762
Sequence: 1 gagaaacgagctactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl :

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3762	100.0	3762	6	AX505122	AX505122 Sequence
2	3596.6	95.6	4032	6	AK126682	AK126682 Homo sapi
3	3590	95.4	4136	6	CQ841482	CQ841482 Sequence
4	3590	95.4	4136	9	AK124520	AK124520 Homo sapi
5	3522.2	93.6	185467	9	AL606804	AL606804 Human DNA
6	3511.6	93.3	141268	2	AC074365	AC074365 Homo sapi
7	2604.8	69.2	2648	9	BC024174	BC024174 Homo sapi
8	1673.4	44.5	1709	6	CQ842314	CQ842314 Sequence
9	1673.4	44.5	1709	9	AK123798	AK123798 Homo sapi
10	1504.6	40.0	1924	9	AB060908	AB060908 Macaca fa
11	1068	28.4	1128	6	CQ488422	CQ488422 Sequence
12	1068	28.4	1128	6	CQ488563	CQ488563 Sequence
13	1068	28.4	1128	6	CQ488741	CQ488741 Sequence
14	1068	28.4	1128	6	CQ488930	CQ488930 Sequence
15	1068	28.4	1128	6	CQ494252	CQ494252 Sequence
16	1068	28.4	1128	6	CQ494397	CQ494397 Sequence
17	1068	28.4	1128	6	CQ494580	CQ494580 Sequence
18	1068	28.4	1128	6	CQ494775	CQ494775 Sequence
19	994.8	26.4	121353	2	AC023386	AC023386 Homo sapi

C	20	697.8	18.5	121353	2	AC023386	AC023386 Homo sapi
	21	394.2	10.5	433	6	CQ469395	CQ469395 Sequence
	22	393.8	10.5	397	6	CQ478563	CQ478563 Sequence
	23	392.2	10.4	421	6	CQ499427	CQ499427 Sequence
	24	392.2	10.4	437	6	CQ499760	CQ499760 Sequence
	25	392.2	10.4	437	6	CQ508397	CQ508397 Sequence
	26	392.2	10.4	437	6	CQ508729	CQ508729 Sequence
	27	391.8	10.4	404	6	CQ478228	CQ478228 Sequence
	28	375	10.0	446	6	CQ469060	CQ469060 Sequence
C	29	354	9.4	70311	9	AC002390	AC002390 Human DNA
	30	353.6	9.4	174875	2	AC024682	AC024682 Homo sapi
	31	342.8	9.1	359	6	CQ477717	CQ477717 Sequence
	32	342.2	9.1	400	6	CQ468549	CQ468549 Sequence
	33	342.2	9.1	401	6	CQ498920	CQ498920 Sequence
	34	342.2	9.1	401	6	CQ507889	CQ507889 Sequence
	35	342.2	9.1	401	6	CQ508110	CQ508110 Sequence
C	36	340.6	9.1	149490	9	AL355344	AL355344 Human DNA
	37	340.6	9.1	168312	2	AC007484	AC007484 Homo sapi
C	38	340.6	9.1	191161	2	AC016043	AC016043 Homo sapi
	39	300.8	8.0	150424	9	AL391811	AL391811 Human DNA
C	40	300.8	8.0	190793	2	AC034298	AC034298 Homo sapi
	41	300.8	8.0	199594	9	AC093560	AC093560 Homo sapi
	42	295.8	7.9	168055	9	AC090796	AC090796 Homo sapi
	43	295.8	7.9	191037	9	AC091561	AC091561 Homo sapi
C	44	295.8	7.9	191234	2	AC018681	AC018681 Homo sapi
	45	277	7.4	98697	9	AC004854	AC004854 Homo sapi

ALIGNMENTS

RESULT 1
AX505122
LOCUS AX505122 3762 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1 from Patent WO0246389.
ACCESSION AX505122
VERSION AX505122.1 GI:23386429
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Nocka,K., Pirozzi,G. and Einstein,R.
TITLE Novel genes associated with allergic hypersensitivity and mast cell activation
JOURNAL Patent: WO 0246389-A 1 13-JUN-2002;
FEATURES
source UCB, S.A. (BE)
location/Qualifiers
1..3762
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
25..432
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD48808.1"
/db_xref="GI:23386430"
/translation="MGNYLLRLKSLGKQKPKKGNPDEERKROEMTTFERKLODQD
KKQSVSTNOENENGSGSEVCYIVNHIFHQSSLSNSDDGVDGNIDSLRKRQF
RRSETEYALLTSVSRPCSTHEHDYEVVPH"

Query Match 100.0%; Score 3762; DB 6; Length 3762;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGAAACCGAGTCACGTGTAAGATGGGAATATCTCTCGGAAAACCTCAGTTCCTG 60
Db 1 GAGAAACCGAGTCACGTGTAAGATGGGAATATCTCTCGGAAAACCTCAGTTCCTG 60
QY 61 GGAGAGAACCAAGAGCCAGAACCCAGAACCCAGATCAGGAAGAAACCGAGGAA 120

This Page Blank (uspto)

QY 2287 TCAGTGCTCAAAAGCTTTCAGATGTTAAGCTGGTATGCAGTTCATGCCAGTGATCCGAG 2346
Db 10164 TCCCTTGATTTAAAAGCTCAGATTCAGGGCTGAGTGTGGTGTCTCACACCTGTAATCCAG 10223

Search completed: November 19, 2004, 01:44:09
Job time : 1747 secs

```
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-03-02
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1384
; LENGTH: 95960
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(95960)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1384

Query Match          5.3%; Score 201.2; DB 13; Length 95960;
Best Local Similarity 72.2%; Pred. No. 1.5e-28;
Matches 322; Conservative 0; Mismatches 108; Indels 16; Gaps 4;

Qy 954 TGTAGATCACACACTGTCTGAGGATTCATCTATGAGACTTTGCTACATAACAGAGAC 1013
Db 4743 TTTACTGCTATTCTCTCTGAGGCTGATGCTGTGAAACTTTGCTCTTCAACAGAAC 4802

Qy 1014 CTTGGTTTCCACACC-CCITTAATTTAGCTAAAGCAATCTTTCTACTGACTTCTTAAG 1072
Db 4803 CTTGGTCTCTATAGCGTCCCTCATATTAACCTGAAACAATTTCTTTCTATTGACTTC--AAG 4860

Qy 1073 TCCTTAGACAAAGCTTAACTCTTTCAACCAATTGCCAATCAGACAAACTTTGMAATCTACC 1132
Db 4861 TCCTTAGACACAGTTGAATCTTTTAAACCAATTGCCAATCAGAAATCTTAGACTCCACC 4920

Qy 1133 TATGACCTGTAAGCTCTCTCTGCTTCAAGATCTTGCTCTTTAAAGCTGAACCGATGTC 1192
Db 4921 TAAGACCTGTAA--CCACCCCTCAAGGTATCTTGCTCTTTAGGCTGAATCAATGTAT 4978

Qy 1193 ACTTTCCATTTAATGATTTATGCTTTGCTTTGTAACCTCTCCCTTAAATGTATAAA 1252
Db 4979 ACCTTCCATGTTGACTTATGATTTTACCTAAATTCCTGTCTCCCTTAAATGTACAAA 5038

Qy 1253 AGTAAACGGTGACCTGACACCTCAGGCACACTTTCTCAGACCTCTCTGAGAGTATCC 1312
Db 5039 ACCAATGATACCTTGACCGCTTGCGCACACTTTCTCAGATCTCTTGAGACTGTCCCC 5098

Qy 1313 CAGCCCATGGTAAAGT-----CATGTGGCTCAGAATCAACCTCTTTAAATATTT 1361
Db 5099 CAGCCCATGGTCACTCATATGCGACCATATGGGCTCAGAAATAAACCTCTTTAAATATTT 5158

Qy 1362 TACAGATTTGGGTTTGGTTACCAA 1387
Db 5159 TACAGACGAGTAATTTGGGAGGCTAA 5184

RESULT 12
US-10-087-192-1390
; Sequence 1390, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1372
```

```
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1390
; LENGTH: 95960
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(95960)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1390

Query Match          5.3%; Score 201.2; DB 13; Length 95960;
Best Local Similarity 72.2%; Pred. No. 1.5e-28;
Matches 322; Conservative 0; Mismatches 108; Indels 16; Gaps 4;

Qy 954 TGTAGATCACACACTGTCTGAGGATTCATCTATGAGACTTTGCTACATAACAGAGAC 1013
Db 4743 TTTACTGCTATTCTCTCTGAGGCTGATGCTGTGAAACTTTGCTCTTCAACAGAAC 4802

Qy 1014 CTTGGTTTCCACACC-CCITTAATTTAGCTAAAGCAATCTTTCTACTGACTTCTTAAG 1072
Db 4803 CTTGGTCTCTATAGCGTCCCTCATATTAACCTGAAACAATTTCTTTCTATTGACTTC--AAG 4860

Qy 1073 TCCTTAGACAAAGCTTAACTCTTTCAACCAATTGCCAATCAGACAAACTTTGMAATCTACC 1132
Db 4861 TCCTTAGACACAGTTGAATCTTTTAAACCAATTGCCAATCAGAAATCTTAGACTCCACC 4920

Qy 1133 TATGACCTGTAAGCTCTCTCTGCTTCAAGATCTTGCTCTTTAAAGCTGAACCGATGTC 1192
Db 4921 TAAGACCTGTAA--CCACCCCTCAAGGTATCTTGCTCTTTAGGCTGAATCAATGTAT 4978

Qy 1193 ACTTTCCATTTAATGATTTATGCTTTGCTTTGTAACCTCTCCCTTAAATGTATAAA 1252
Db 4979 ACCTTCCATGTTGACTTATGATTTTACCTAAATTCCTGTCTCCCTTAAATGTACAAA 5038

Qy 1253 AGTAAACGGTGACCTGACACCTCAGGCACACTTTCTCAGACCTCTCTGAGAGTATCC 1312
Db 5039 ACCAATGATACCTTGACCGCTTGCGCACACTTTCTCAGATCTCTTGAGACTGTCCCC 5098

Qy 1313 CAGCCCATGGTAAAGT-----CATGTGGCTCAGAATCAACCTCTTTAAATATTT 1361
Db 5099 CAGCCCATGGTCACTCATATGCGACCATATGGGCTCAGAAATAAACCTCTTTAAATATTT 5158

Qy 1362 TACAGATTTGGGTTTGGTTACCAA 1387
Db 5159 TACAGACGAGTAATTTGGGAGGCTAA 5184

RESULT 13
US-10-087-192-1372/c
; Sequence 1372, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1372
```



```
RESULT 7
US-10-027-632-31701/c
; Sequence 31701, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31701
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-31701

Query Match      5.4%; Score 205; DB 13; Length 810;
Best Local Similarity 73.5%; Pred. No. 3.4e-30;
Matches 302; Conservative 0; Mismatches 105; Indels 4; Gaps 3;

Qy 967 CTGCTGAGGATTCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGGTTTCCACA 1026
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
778 CTCTTGAAGGTTGCAACCTGGAGGGCTCTTTTACAATAACAAGAAATCTGGGCTTCCAC 719
Qy 1027 ACCCTTTATTTAGCTAAA-GCAATCTTTTCTACTGACTTTTAAAGTCTTTAGCAAAAG 1085
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
718 TCTCCCTATCTTAAATAAACCTTTCTTTCTACTGACTTC--AAGTCITTTACAAAG 661
Qy 1086 CTTAACTCTTTCAACCAATGCCAATCAGACAACTTTGAATCTACCTATGACCTGTAAG 1145
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
660 CTTAACTCTTTCAACCAATGCCAATGCCCCAACAATAATCTTTGAATCCACCT-CCACCTGCCAA 602
Qy 1146 CTCTCTCTGCTTCAAGATCTTGCTCTTTAAGCTGAACCGATGTGCACCTTCCATTAA 1205
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
601 GGCCCCCACTTTGAGATGTCTCTCTCTTCAGGCTGAACCAATGTATATCTTACATGTAT 542
Qy 1206 TGATTTATGCTTTGCTTTGTAACCTCTCTCCCTAAAATGTATAAAAGTAAACGGTGAC 1265
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
541 TGATTTATGCTTTGCTTTTAACTCTATCTCCCTAAAATGTATAAAACCAATCTCTTAAC 482
Qy 1266 CTGACCACTCAGGCACACTTTTCTCAGGACCTCTCAGAGTGTATCCAGGCCATGGTAA 1325
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 CCAACTACCTCAGGCACACTTTCTCAGGACTCTTGTAGACTGTTCCCTGGGCCATGGTCA 422
Qy 1326 GTCATGTGGCTCAGAACTCAACCTCTTAAATATTTTACAGAATTTGGGTT 1376
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 CTCATATGGCTCTGAATAAACCTCTTTAGATATTTTACAGAGTTTACTTT 371
```

```
RESULT 8
US-10-027-632-31701/c
; Sequence 31701, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31701
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-31701
```

```
Query Match      5.4%; Score 205; DB 15; Length 810;
Best Local Similarity 73.5%; Pred. No. 3.4e-30;
Matches 302; Conservative 0; Mismatches 105; Indels 4; Gaps 3;

Qy 967 CTGCTGAGGATTCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGGTTTCCACA 1026
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
778 CTCTTGAAGGTTGCAACCTGGAGGGCTCTTTTACAATAACAAGAAATCTGGGCTTCCAC 719
Qy 1027 ACCCTTTATTTAGCTAAA-GCAATCTTTTCTACTGACTTTTAAAGTCTTTAGCAAAAG 1085
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
718 TCTCCCTATCTTAAATAAACCTTTCTTTCTACTGACTTC--AAGTCITTTACAAAG 661
Qy 1086 CTTAACTCTTTCAACCAATGCCAATCAGACAACTTTGAATCTACCTATGACCTGTAAG 1145
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
660 CTTAACTCTTTCAACCAATGCCAATGCCCCAACAATAATCTTTGAATCCACCT-CCACCTGCCAA 602
Qy 1146 CTCTCTCTGCTTCAAGATCTTGCTCTTTAAGCTGAACCGATGTGCACCTTCCATTAA 1205
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
601 GGCCCCCACTTTGAGATGTCTCTCTCTTCAGGCTGAACCAATGTATATCTTACATGTAT 542
Qy 1206 TGATTTATGCTTTGCTTTGTAACCTCTCTCCCTAAAATGTATAAAAGTAAACGGTGAC 1265
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
541 TGATTTATGCTTTGCTTTTAACTCTATCTCCCTAAAATGTATAAAACCAATCTCTTAAC 482
Qy 1266 CTGACCACTCAGGCACACTTTTCTCAGGACCTCTCAGAGTGTATCCAGGCCATGGTAA 1325
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 CCAACTACCTCAGGCACACTTTCTCAGGACTCTTGTAGACTGTTCCCTGGGCCATGGTCA 422
Qy 1326 GTCATGTGGCTCAGAACTCAACCTCTTAAATATTTTACAGAATTTGGGTT 1376
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 CTCATATGGCTCTGAATAAACCTCTTTAGATATTTTACAGAGTTTACTTT 371
```

```
RESULT 9
US-10-428-487-11/c
; Sequence 11, Application US/10428487
; Publication No. US20040006780A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA K.
; APPLICANT: GERBER, HANS-PETER
; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
; FILE REFERENCE: 09800080-0103
; CURRENT APPLICATION NUMBER: US/10/428,487
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/815,153
; PRIOR FILING DATE: 2001-03-21
```



```

Db 187575 TTTCAAAATAGAGTGCACTATTGCTGTAATAAAAAAAGTTGAAGGGCTGGATACGGTGGCTC 187516
Qy 2331 ATGCCAGTGATCCGAGTACTTTTGGGAAGCCAGAGCAGGTGGATCTCTTTGAGCCGAGGAGT 2390
Db 187515 ATGCTGTAAATCCAGTACTTTTGGGAGCCGAGGAGCAGATTGCCCTGAGGTGAGGAGT 187456
Qy 2391 TTGAGGCAGACTGCAACAACAGTGAACCTCGTTTCTCAAAATAATTAATAAATTTAGC 2450
Db 187455 TTGAGACAGAGCTGTCAACATGTGTGAACCCCATCTCCACTAAAAATACAAAAATTTAGC 187396
Qy 2451 CAGGTGTGTGTGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGGATG 2510
Db 187395 CAGGCATGTGTGTGTGCTGCTTAACTCTACTCTCAGGAGGTTGAGGAGGAGGATCA 187336
Qy 2511 TTGAGACTGGGAGGTGAGGCTGAACCTGAGCTGAGCCAGGATCTTGGCCACCACTTCAGGCTTG 2570
Db 187335 CTTGAACCCAGGAGAGCGAGGTGTCAGTGAAGCCGACAGTGCCACTGCCTCCAGGCTG 187276
Qy 2571 GGCACAGAGTGAGACCCCTGTCTCAAAAAAATAAAAAA 2608
Db 187275 GGCACAGAGTCACTGTCTCAAAAAAATAAAAAAAGA 187238

```

```

RESULT 3
US-10-027-632-125465/c
; Sequence 125465, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.123
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 125465
; LENGTH: 830
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-125465

```

```

Query Match 6.0%; Score 224.6; DB 13; Length 830;
Best Local Similarity 68.9%; Pred. No. 4.9e-34;
Matches 399; Conservative 0; Mismatches 164; Indels 16; Gaps 6;

Qy 2086 GATTTCGAATATTTCCATACATATAATGAGAGAGTTGGAAATGGGATTCAGTCTAAT 2145
Db 812 GATTTCGAATATTTCCATACATATAATGAGAGAGTTGGAAATGGGATTCAGTCTAAT 2145
Qy 2146 CATAAAATTCACCTATGTTTGCATATACACCTTATCTGAATAGCCTGGAAGTAAATTTATA 2205
Db 752 CAAGAGCTCA -TTTGTTCATATACACCTTATCTGAATAGCCTGGAAGTAAATTTATA 2205
Qy 2206 CAATATTTAAATATTTTATGCTTGAACAGAGTTTGGCAGATTCGACATCAGAAAG 2265
Db 694 CAGTGCCTTTTATA -TTTCTGTGTGAACAAAGTTTGGTACATGTAACCAT -GGAAAG 638

```

```

Qy 2266 CAGAAGTGTCACTATTTCAAGTCAGTCTCAAAAAGTTTCAGATGTTAAGCTGTGTATGC 2325
Db 637 CAAAGGTGTAC -GCTGGTATTGAAAAAGTTTCAGATGTTAAGCTGTGTATGC 586
Qy 2326 AGTTTCATGCCAGTATCCGAGTACTTTTGGGAAGCCAGAGCAGGTGGATCTCTTTGAGCCCA 2385
Db 585 GGCTCAAGCCAGTAATCCAGCAGTGTAGGAGGCTGAGGAGGAGGATCCCTTGGAGCCCA 526
Qy 2386 GGAGTTTGGAGCCAGACTGCACAAACAGTGAACCTCGTTTCTTCAAAATAATTTAAAAA 2445
Db 525 GGAGTTTGGAGCCAGCTGGGCAACATGGCAAAACCTCTCTCTTACAAAAAT -ACAAG 468
Qy 2446 TTAGCCAGGTGTGTGTGCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAG 2505
Db 467 TTTGCCAGGTGTGTGTAGCAAGTGCCTGTAGTCTCAGGTCTTCAAGGCGGCTGGGGTGGAG 408
Qy 2506 GATTGTTTGAAGCTGGGAGGTGAGGCTGAACCTGAGCAGGATCTTCCACCAACATTTCCA 2565
Db 407 GATTGTTTGAAGCCAGCAAGTGGAGGCTGAGTGAAGCGAGATGGCAGACTGCTCTCCG 348
Qy 2566 GCTTGGCAACAGAGTGAACCTGTCTC -AAAAAATAAAGTTTCAGATTTTGGAG 2624
Db 347 GCCTGGCGACAGAGTAAGACCCCTGTCTCTCCACACACACAAATTCAGATTTTGCAG 288
Qy 2625 CATTCAGATCTTCAGATTAGGAGTTTCAACCTGTACT 2663
Db 287 CGTTTCAGAAATTTTGGATTAGGCTGCTGAACCTGTAAAT 249

```

```

RESULT 4
US-10-027-632-125465/c
; Sequence 125465, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 125465
; LENGTH: 830
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-125465

```

```

Query Match 6.0%; Score 224.6; DB 15; Length 830;
Best Local Similarity 68.9%; Pred. No. 4.9e-34;
Matches 399; Conservative 0; Mismatches 164; Indels 16; Gaps 6;

Qy 2086 GATTTCGAATATTTCCATACATATAATGAGAGAGTTGGAAATGGGATTCAGTCTAAT 2145
Db 812 GATTTCGAATATTTCCATACATATAATGAGAGAGTTTGGGATGTGACATAGGCTCAAA 753
Qy 2146 CATAAAATTCACCTATGTTTGCATATACACCTTATCTGAATAGCCTGGAAGTAAATTTATA 2205
Db 752 CAAGAGCTCA -TTTGTTCATATACACCTTATTCATAGCTGGAAGTAAATTTATA 2205

```


[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2004, 18:25:09 ; Search time 1741 Seconds
(without alignments)
11676.296 Million cell updates/sec

Title: US-10-005-907-1
Perfect score: 3762
Sequence: 1 gagaaaccgagtcactgtga.....aaaaaaaaaaaaaaaa 3762

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3627888 seqs, 270181610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3762	100.0	3762	15	US-10-005-907-1
C 2	234.4	6.2	191584	17	Sequence 1, Appli
C 3	224.6	6.0	830	13	Sequence 800, App
C 4	224.6	6.0	830	15	Sequence 125465,
C 5	207.4	5.5	58320	17	Sequence 125465,
C 6	207.4	5.5	63313	13	Sequence 88, Appl
C 7	205	5.4	810	13	Sequence 526, App
C 8	205	5.4	810	15	Sequence 31701, A
C 9	203.8	5.4	301692	16	Sequence 31701, A
C 10	203.8	5.4	310268	17	Sequence 11, Appl
C 11	201.2	5.3	95960	13	Sequence 195, App
C 12	201.2	5.3	95960	13	Sequence 1384, Ap
					Sequence 1390, Ap

C 13	201.2	5.3	130207	13	US-10-087-192-1372	Sequence 1372, Ap
C 14	200.4	5.3	17239	9	US-09-764-869-2326	Sequence 2326, Ap
C 15	200.4	5.3	17239	14	US-10-091-504-2326	Sequence 2326, Ap
C 16	200.4	5.3	17239	16	US-10-227-577-2326	Sequence 2326, Ap
C 17	199.8	5.3	181259	16	US-10-456-930-2	Sequence 2, Appli
C 18	199.6	5.3	73145	13	US-10-087-192-274	Sequence 274, App
C 19	199.2	5.3	728	13	US-10-027-632-19728	Sequence 19728, A
C 20	199.2	5.3	728	15	US-10-027-632-19728	Sequence 2, Appli
C 21	197.4	5.2	290	10	US-09-854-867-2	Sequence 6056, App
C 22	197.4	5.2	1743	15	US-10-091-414-236	Sequence 236, App
C 23	197.4	5.2	1746	10	US-09-764-891-6057	Sequence 6057, App
C 24	197.4	5.2	1746	15	US-10-091-414-237	Sequence 237, App
C 25	197.4	5.2	1746	15	US-10-091-414-237	Sequence 86, Appl
C 26	197.4	5.2	6686	9	US-09-736-960-86	Sequence 46, Appl
C 27	197	5.2	166043	16	US-10-235-192A-46	Sequence 8396, App
C 28	196.8	5.2	27154	10	US-09-764-891-8396	Sequence 169, App
C 29	196.2	5.2	11750	17	US-10-367-094-189	Sequence 105966,
C 30	196	5.2	586	13	US-10-027-632-105966	Sequence 105966,
C 31	196	5.2	586	15	US-10-027-632-105966	Sequence 1, Appli
C 32	195.4	5.2	288	14	US-10-115-278-1	Sequence 1486, App
C 33	194.8	5.2	54169	13	US-10-087-192-1486	Sequence 2670, App
C 34	193.8	5.2	8207	9	US-09-764-877-2670	Sequence 741, App
C 35	193.8	5.2	8207	10	US-09-764-872-741	Sequence 2670, App
C 36	193.8	5.1	598	13	US-10-242-515-2670	Sequence 269156,
C 37	193.6	5.1	598	15	US-10-027-632-269156	Sequence 269156,
C 38	193.6	5.1	598	15	US-10-027-632-269156	Sequence 29373, A
C 39	193.6	5.1	628	13	US-10-027-632-29373	Sequence 29373, A
C 40	193.6	5.1	628	15	US-10-027-632-29373	Sequence 265848,
C 41	193.2	5.1	1599	13	US-10-027-632-265848	Sequence 265848,
C 42	193.2	5.1	1599	15	US-10-027-632-265848	Sequence 76, Appl
C 43	193.2	5.1	96588	15	US-10-085-117-76	Sequence 11, Appl
C 44	193.2	5.1	185695	14	US-10-020-141-11	Sequence 1, Appli
C 45	193.2	5.1	185695	14	US-10-017-721-1	

ALIGNMENTS

RESULT 1
US-10-005-907-1
; Sequence 1, Application US/10005907
; Publication No. US20030166881A1
; GENERAL INFORMATION:
; APPLICANT: Union Chimique Belge, S.A.
; APPLICANT: No. US20030166881A1, Karl
; APPLICANT: Pirozzi, Gregory
; APPLICANT: Einstein, Richard
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAST
; FILE REFERENCE: 053529-5005
; CURRENT APPLICATION NUMBER: US/10/005,907
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 3762
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(432)
; OTHER INFORMATION:
US-10-005-907-1

Query Match 100.0%; Score 3762; DB 15; Length 3762;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGAAACCGAGTCAGTCGTGAAAAGATGGGAAATATCTCTCGGAAAACCTCAGTTGCGCTG 60
Db 1 GAGAAACCGAGTCAGTCGTGAAAAGATGGGAAATATCTCTCGGAAAACCTCAGTTGCGCTG 60
QY 61 GGAGAGATCAAAAGAACGCCCAAGAAAGAAACCCAGATGAGGAAAGAAACCGGAGGAA 120

Page Blank (uspto)

Query Match 10.4%; Score 392.2; DB 5; Length 437;
 Best Local Similarity 99.2%; Pred. No. 1.5e-47;
 Matches 394; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1571	CTATGTACTGGATATTTTTCATATATAAACTTGCAGTAATAGTTCAAAAATTAATAGTTTT	1630
Db	41	CGAGGTACTGGATATTTTTCATATATAAACTTGCAGTAATAGTTCAAAAATTAATAGTTTT	100
QY	1631	TGACATTTGGCTTTTCTGAGAGAGAAATTCGAAGTCTCACAATTAATAAAGATGAAAT	1690
Db	101	TGACATTTGGCTTTTCTGAGAGAGAAATTCGAAGTCTCACAATTAATAAAGATGAAAT	160
QY	1691	GAAGCATATATAATTTGTCAATTTTTCATTTTCTAGTCAACAGAGATCGAAGATTCT	1750
Db	161	GAAGCATATATAATTTGTCAATTTTTCATTTTCTAGTCAACAGAGATCGAAGATTCT	220
QY	1751	GTTCAAAATATTAGTAAATAATTGAAATAAATCTTGTGCTTATATTTTGTTCACACACT	1810
Db	221	GTTCAAAATATTAGTAAATAATTGAAATAAATCTTGTGCTTATATTTTGTTCACACACT	280
QY	1811	AGTTAATTTAACTGTGACTAGTTATCTCTACCGAAGGTGGATGTGTAGTTCTGTGTTTT	1870
Db	281	AGTTAATTTAACTGTGACTAGTTATCTCTACCGAAGGTGGATGTGTAGTTCTGTGTTTT	340
QY	1871	AAAAATCAAGCAAACTGGAAAATAATCCATCTAATTATGCTTTCTTTCCCAAGAGTTTT	1930
Db	341	AAAAATCAAGCAAACTGGAAAATAATCCATCTAATTATGCTTTCTTTCCCAAGAGTTTT	400
QY	1931	TTAATGATATGCCAGCTTCCTAATTTTGGAGACAAAAG	1967
Db	401	TTAATGATATGCCAGCTTCCTAATTTTGGAGACAAAAG	437

Search completed: November 18, 2004, 14:08:43
 Job time : 1658 secs

```
Db 376 TTAATGATATGCCAGCTTCTTAATTTGGAGACAAAAGCCTTAATTG 421

RESULT 14
ID ABV31609 standard; cDNA; 437 BP.
XX AC ABV31609;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 31600.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PA Schlegel R, Endege WO, Monahan JE;
XX PI WPI; 2001-662795/76.
XX DR Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 6790; 11750pp; English.
XX SQ Sequence 437 BP; 150 A; 63 C; 79 G; 145 T; 0 U; 0 Other;

Query Match 10.4%; Score 392.2; DB 5; Length 437;
Best Local Similarity 99.2%; Pred. No. 1.5e-47;
Matches 394; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1571 CTATGCTAGTATTTTGGCATATAAACTTGCAGTAAATAGTTCAAAATTAATAGTTTT 1630
Db 41 CGAGTACTGGATATTTTGGCATATAAACTTGCAGTAAATAGTTCAAAATTAATAGTTTT 100
QY 1631 TGACATTGGCTTTCTGAGAGAGAAATTTGAAGTGCACAAAATAAAAAGATGAAAT 1690
Db 101 TGACATTGGCTTTCTGAGAGAGAAATTTGAAGTGCACAAAATAAAAAGATGAAAT 160
QY 1691 GAAGCATATATAATTTGCAATTTTTCATTTTCTAGTCAACAGAGAAATCGAAGATTCT 1750
Db 161 GAAGCATATATAATTTGCAATTTTTCATTTTCTAGTCAACAGAGAAATCGAAGATTCT 220

1751 GTTCAAAATATTAGTAAAAATTGAAATAAACTTGTGCTTAATATTTGTTTGGCAACACT 1810
221 GTTCAAAATATTAGTAAAAATTGAAATAAACTTGTGCTTAATATTTGTTTGGCAACACT 280
1811 AGTTAAATTTAACTGTGACTAGTTATCTCTACCGAAGGTGGATGTAGTTCTGGTTTT 1870
281 AGTTAAATTTAACTGTGACTAGTTATCTCTACCGAAGGTGGATGTAGTTCTGGTTTT 340
1871 AAAATTCAGCAAACTGGAAAAATAATCCATCTAATTATGCTTTTCCCAAGAAGTTTT 1930
341 AAAATTCAGCAAACTGGAAAAATAATCCATCTAATTATGCTTTTCCCAAGAAGTTTT 400
1931 TTAATGATATGCCAGCTTCTCAATTTGGAGACAAAAG 1967
401 TTAATGATATGCCAGCTTCTCAATTTGGAGACAAAAG 437

RESULT 15
ID ABV40245 standard; cDNA; 437 BP.
XX AC ABV40245;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 40236.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PA Schlegel R, Endege WO, Monahan JE;
XX PI WPI; 2001-662795/76.
XX DR Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 8129; 11750pp; English.
XX SQ The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX SQ Sequence 437 BP; 150 A; 63 C; 79 G; 145 T; 0 U; 0 Other;
```

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 XX Claim 1; Page 1680; 11750pp; English.
 XX
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 XX Sequence 397 BP; 142 A; 53 C; 64 G; 138 T; 0 U; 0 Other;
 XX
 XX Query Match 10.5%; Score 393.8; DB 5; Length 397;
 XX Best Local Similarity 99.5%; Pred. No. 8.8e-48;
 XX Matches 395; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1573 ATGTACTGGATATTTTGCATATAAACTTGCAGTAATAGTCAAAAATTAATAGTTTGG 1632
 Db 1 AGGTACTGGATATTTTGCATATAAACTTGCAGTAATAGTCAAAAATTAATAGTTTGG 60
 QY 1633 ACATTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATTAATAAGATGA 1692
 Db 61 ACATTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATTAATAAGATGA 120
 QY 1693 AGCATATATAATGTCAAATTTTCAATTTTCTAGTCAACAGAGATCGAAGATTCGT 1752
 Db 121 AGCATATATAATGTCAAATTTTCAATTTTCTAGTCAACAGAGATCGAAGATTCGT 180
 QY 1753 TCAATATATTAGTAAATAATGAAATAAATCTGTGCTTATATTTGTTCCACACACTAG 1812
 Db 181 TCAATATATTAGTAAATAATGAAATAAATCTGTGCTTATATTTGTTCCACACACTAG 240
 QY 1813 TTAATTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGTAGTTTCTGGTTTAA 1872
 Db 241 TTAATTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGTAGTTTCTGGTTTAA 300
 QY 1873 AATTCAGCAAACTGGAATAATTCATCTAATATGCTTTCTTCCCAAGAGTTTCTT 1932
 Db 301 AATTCAGCAAACTGGAATAATTCATCTAATATGCTTTCTTCCCAAGAGTTTCTT 360
 QY 1933 AATGATATCCAGCTTCCCTAAATTTGGAGACAAAAGCC 1969
 Db 361 AATGATATCCAGCTTCCCTAAATTTGGAGACAAAAGCC 397
 XX
 XX RESULT 13
 XX ABV31276
 ID ABV31276 standard; cDNA; 421 BP.
 XX
 XX AC ABV31276;
 XX
 XX 16-SEP-2002 (first entry)
 DT
 XX Human prostate expression marker cDNA 31267.
 DE
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX pharmacogenomic marker; gene; ss.
 KW

XX Homo sapiens.
 OS
 XX WO200160860-A2.
 PN
 XX 23-AUG-2001.
 PD
 XX
 XX 20-FEB-2001; 2001WO-US005171.
 PF
 XX 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 PR
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA
 XX Schlegel R, Endege WO, Monahan JE;
 PI
 XX WPI; 2001-662795/76.
 DR
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 XX Claim 1; Page 6740; 11750pp; English.
 PS
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 XX Sequence 421 BP; 143 A; 61 C; 71 G; 145 T; 0 U; 1 Other;
 SQ
 XX Query Match 10.4%; Score 392.2; DB 5; Length 421;
 XX Best Local Similarity 97.8%; Pred. No. 1.5e-47;
 XX Matches 397; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1571 CTATGTACTGGATATTTTGCATATAAACTTGCAGTAATAGTCAAAAATTAATAGTTT 1630
 Db 16 CGAGGTACTTTATATTTTGCATATAAACTTGCAGTAATAGTCAAAAATTAATAGTTT 75
 QY 1631 TGACATTTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATTAATAAGATGAAT 1590
 Db 76 TGACATTTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATTAATAAGATGAAT 135
 QY 1691 GAAGCATATATAATTTGTCATATTTTCAATTTTCTAGTCAACAGAGATCGAGGATCT 1750
 Db 136 GAAGCATATATAATTTGTCATATTTTCAATTTTCTAGTCAACAGAGATCGAGGATCT 195
 QY 1751 GTTCAAAATATTAGTAAAAATTTGAAAATTAACATTTGCTTATATTTTGTTCGAACACAT 1810
 Db 196 GTTCAAAATATTAGTAAAAATTTGAAAATTAACATTTGCTTATATTTTGTTCGAACACAT 255
 QY 1811 AGTTAATTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGTAGTTTCTGGTTT 1870
 Db 256 AGTCAATTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGTAGTTTCTGGTTT 315
 QY 1871 AAAATTTCAAGCAAACTGGAAAATTAATCCATCTAATTTATGCTTTCTTCCCAAGAGTTT 1930
 Db 316 AAAATTTCAAGCAAACTGGAAAATTAATCCATCTAATTTATGCTTTCTTCCCAAGAGTTT 375
 QY 1931 TTAATGATATGCCAGCTTCCCTAATTTGGAGACAAAAGCCTTAATTG 1976

Db 222 GGAGAGAAATCAAAGAGCCCAAGAAAGGAAGAAACCCAGATGAGGAGAAAGAAACGGCAGGAA 281
 Qy 121 ATGACTACATTTGAAGAAAGAACTTCAAGATCAAGATAAGAAAGCAAGAAAGTTTCATCC 180
 Db 282 ATGACTACATTTGAAGAAAGAACTTCAAGATCAAGATAAGAAAGCAAGAAAGTTTCATCC 341
 Qy 181 ACTTCTAATCAGGAAAGCAAGATGGCAGTGGTCTGGAAGAGTGTGCTACACTGTCAAT 240
 Db 342 ACTTCTAATCAGGAAAGCAAGATGGCAGTGGTCTGGAAGAGTGTGCTACACTGTCAAT 401
 Qy 241 AATCACAATCCCAATCAGAGATCCCTCCCTGAGCTCCAATGATGATGGCTATGAGAACATT 300
 Db 402 AATCACAATCCCAATCAGAGATCCCTCCCTGAGCTCCAATGATGATGGCTATGAGAACATT 461
 Qy 301 GACTCCCTCAAGAAAGTGAAGAGTGTGAGAAAGGTGAGAGACAGATATGCCCTT 360
 Db 462 GACTCCCTCAAGAAAGTGAAGAGTGTGAGAAAGGTGAGAGACAGATATGCCCTT 521
 Qy 361 CTTAGGACTTCTGTTAGTAGGCTTGTTCCTGCAACCCATGAGCATGATTATGAAGTTGTG 420
 Db 522 CTTAGGACTTCTGTTAGTAGGCTTGTTCCTGCAACCCATGAGCATGATTATGAAGTTGTG 581
 Qy 421 TTTCACACTAAATCTCAAGCTGCTTTATCACCTTCCAGCAATGAGACAAATGCAGAA 480
 Db 582 TTTCACACTAAATCTCAAGCTGCTTTATCACCTTCCAGCAATGAGACAAATGCAGAA 641
 Qy 481 TAGCAGACTCTGGCGAAGTTGTTTAC 506
 Db 642 TAGCAGACTCTGGCGAAGTTGTTTAC 667

RESULT 11

ABV01270
 ID ABV01270 standard; cDNA; 433 BP.
 XX
 AC ABV01270;
 XX

13-SEP-2002 (first entry)

Human prostate expression marker cDNA 1261.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX

WO200160860-A2.

23-AUG-2001.

20-FEB-2001; 2001WO-US005171.

17-FEB-2000; 2000US-0183319P.

16-MAR-2000; 2000US-0189862P.

25-MAY-2000; 2000US-0207454P.

09-JUN-2000; 2000US-0211314P.

18-JUL-2000; 2000US-0219007P.

13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;
 WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of
 prostate cells and correlating with presence of prostate cancer, useful
 for detecting presence of prostate cancer, stage of prostate cancer.

Claim 1; Page 295; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising
 a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 433 BP; 146 A; 64 C; 79 G; 144 T; 0 U; 0 Other;

Query Match 10.5%; Score 394.2; DB 5; Length 433;
 Best Local Similarity 99.2%; Pred. No. 7.6e-48;
 Matches 396; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1571 CTATGTACTCGATATTTTTCATATAAATTCGAGTAAATGCGAGTAAATGATGTTT 1630
 Db 35 CGAGTACTCGATATTTTTCATATAAATTCGAGTAAATGCGAGTAAATGATGTTT 94
 Qy 1631 TGACATTTGGCTTTTTCGAGAGAGAAATTCGAGTAAATGCGAGTAAATGATGTTT 1690
 Db 95 TGACATTTGGCTTTTTCGAGAGAGAAATTCGAGTAAATGCGAGTAAATGATGTTT 154
 Qy 1691 GAAGCATATATAATGTCATATTTTTCATATTTTCTAGTCAACAGAGATCGAAGATTCT 1750
 Db 155 GAAGCATATATAATGTCATATTTTTCATATTTTCTAGTCAACAGAGATCGAAGATTCT 214
 Qy 1751 GTTCAATATAGTAAATTAATGAAATTAATGAAATTAATGAAATTAATGAAATTAAT 1810
 Db 215 GTTCAATATAGTAAATTAATGAAATTAATGAAATTAATGAAATTAATGAAATTAAT 274
 Qy 1811 AGTTAATTTAACCTGTGACTAGTTATCTCTACCGAAGTGGATGTGTAGTTTCTGGTTT 1870
 Db 275 AGTTAATTTAACCTGTGACTAGTTATCTCTACCGAAGTGGATGTGTAGTTTCTGGTTT 334
 Qy 1871 AAAATTCAGCAAACTGGAAAAATTAATCCATCTAATATGCTTTTCCCAAGAGTTT 1930
 Db 335 AAAATTCAGCAAACTGGAAAAATTAATCCATCTAATATGCTTTTCCCAAGAGTTT 394
 Qy 1931 TTAATGATATGCCAGCTTCCCTAATTTGGAGACAAAGGCC 1969
 Db 395 TTAATGATATGCCAGCTTCCCTAATTTGGAGACAAAGGCC 433

RESULT 12

ABV10439
 ID ABV10439 standard; cDNA; 397 BP.
 XX
 AC ABV10439;
 XX

13-SEP-2002 (first entry)

Human prostate expression marker cDNA 10430.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 pharmacogenomic marker; gene; ss.

OS Homo sapiens.

WO200160860-A2.

23-AUG-2001.

20-FEB-2001; 2001WO-US005171.

17-FEB-2000; 2000US-0183319P.

16-MAR-2000; 2000US-0189862P.

25-MAY-2000; 2000US-0207454P.

09-JUN-2000; 2000US-0211314P.

18-JUL-2000; 2000US-0219007P.

13-DEC-2000; 2000US-0255281P.

|||||
303 ATCGAAGGATCTGTTCAAAATATTAGTAAATGAAAATGAAATCTGCTATATTTG 362
1798 TTTGCAACACACTAGTAAATTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGT 1857
363 TTTGCAACACACTAGTAAATTTAACTGTGACTAGTTATCTCTACCGAAGTGGATGT 422
1858 AGTTTCTGTTTAAATTTCAAGCAAACTGGAAAATATCAATCTAAATATGCTTCTTT 1917
423 AGTTTCTGTTTAAATTTCAAGCAAACTGGAAAATATCAATCTAAATATGCTTCTTT 482
1918 CCCAAGAAGTTTAAATGATATGCGAGCTTCTTAATTTGGAGACAAAAGCCTTAATCA 1977
483 CCCAAGAAGTTTAAATGATATGCGAGCTTCTTAATTTGGAGACAAAAGCCTTAATCA 542
1978 CAATGCAATCAATATATATTTTGTATAGTTACAGTATACAGTGTAGTATCCCTTAG 2037
543 CAATGCAATCAATATATATTTTGTATAGTTACAGTATACAGTGTAGTATCCCTTAG 602
2038 ATGAGATGCTTGGGACCAAGTGTGTTTGGATTTGAGATTTATTTTGGATTTGGAATA 2097
603 ATGAGATGCTTGGGACCAAGTGTGTTTGGATTTGAGATTTATTTTGGATTTGGAATA 662
2098 TTTCCATACATATAATGAGAGAGTTGGAAAATGGGATTCAGTCTAATCATAAATTCAC 2157
663 TTTCCATACATATAATGAGAGAGTTGGAAAATGGGATTCAGTCTAATCATAGATGCAC 722
2158 TTATGTTTGGATATACACCTTATCTGAATAGCCTGAAAGTAAATTTTATACAAATTTTAAA 2217
723 TTATGTTTGGATATACACCTTATCTGAATAGCCTGAAAGTAAATTTTATACAAATTTTAAA 782
2218 TAAATTTATGCTGAAACAGAGTTTCGCGACATTTGACCATCAGAAGCAGAAGTGTAC 2277
783 TAAATTTATGCTGAAACAGAGTTTCGCGACATTTGACCATCAGAAGCAGAAGTGTAC 842
2278 TATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATGTTAAGCTGGTGATTCATGCCAG 2337
843 TATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATGTTAAGCTGGTGATTCATGCCAG 902
2338 TGATCCGAGTACTTTGGGAAGCCAGACAGGTGGATCTCTTGAGCCAGGAGTTGAGGC 2397
903 TGATCCGAGTACTTTGGGAAGCCAGACAGGTGGATCTCTTGAGCCAGGAGTTGAGGC 962
2398 CAGACTGCAACACAGTGAAGCTGTTTCTCAAAATATTAATAATTAAGCCAGGTGT 2457
963 CAGACTGCAACACAGTGAAGCTGTTTCTCAAAATATTAATAATTAAGCCAGGTGT 1022
2458 GGTGGTGCACACCTGTAGTCCAGGCTACTCAGGAGGTGAGG-TAGTAGGATTTGTTGAG 2516
1023 GGTGGTGCACACCTGTAGTCCAGGCTACTCAGGAGGTGAGG-TAGTAGGATTTGTTGAG 1082
2517 ACTGGAGGTTGAGGCTGAACCTAGCCAGGATCTTGCCACCA 2560
1083 ACTGGAGGTTGAGGCTGAACCTAGCCAGGATCTTGCCACCA 1126

RESULT 10
ADC30760
ID ADC30760 standard; cdna; 667 BP.

AC ADC30760;

XX 18-DEC-2003 (first entry)

XX Human novel cdna sequence, SEQ ID NO:842.

XX Human; diagnostic; drug screening; forensics; gene mapping;
XX biodiversity assessment; parkinson's disease; Alzheimer's disease;
XX neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
XX ulcers; osteoporosis; autoimmune disease; cancer;
XX molecular weight marker; food supplement; antiparkinsonian; nootropic;
XX neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;
XX antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

gene therapy; chromosome 1; gene; ss.
XX Homo sapiens.
XX WO2003029271-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030474.
XX 24-SEP-2001; 2001US-0324631P.
XX (HYSE-) HYSEQ INC.
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
XX Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
XX P-PSDB; ADC31731.
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
XX treating conditions such as neurodegenerative diseases, anemias, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer.
XX Claim 1; SEQ ID NO 842; 1185pp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-
XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
XX invention also relates to nucleic acid sequences over 99% identical with
XX the novel human cDNAs. The invention additionally encompasses expression
XX vectors and host cells comprising a nucleic acid of the invention; the
XX recombinant production of a polypeptide of the invention; an antibody
XX against a polypeptide of the invention; a method of detecting
XX polynucleotides or polypeptides of the invention; and methods of
XX identifying a compound which binds to a polypeptide of the invention. The
XX invention further discloses methods of preventing, treating or
XX ameliorating a medical condition; kits comprising polynucleotide probes
XX and/or monoclonal antibodies for carrying out the methods of the
XX invention; methods for the identification of compounds that modulate the
XX expression or activity of the polynucleotide and/or polypeptide; and 767
XX contig sequences corresponding to the cDNA sequences of the invention
XX (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
XX -ADC33394). The nucleic acids and polypeptides of the invention are
XX useful in diagnostics, drug screening, forensics, gene mapping, in the
XX identification of mutations responsible for genetic disorders or other
XX traits, for assessing biodiversity, and in producing many other types of
XX data and products dependent on DNA and amino acid sequences. They are
XX also used for treating diseases such as Parkinson's disease, Alzheimer's
XX disease and other neurodegenerative diseases, anaemia, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer. The nucleic acids may also be used as hybridisation probes or
XX primers, and in the recombinant production of a protein. The polypeptides
XX are also useful in generating antibodies, as molecular weight markers,
XX and as food supplements. The present sequence represents a specifically
XX claimed human cDNA sequence of the invention. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 667 BP; 216 A; 147 C; 154 G; 150 T; 0 U; 0 Other;

Query Match 13.5%; Score 506; DB 10; Length 667;
Best Local Similarity 100.0%; Pred. No. 6.8e-64;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAAACCGAGTCACCTGTGAAAAGATGGGAAATATCTCTCGGAAACCTCAGTTGCTG 60
DB 162 GAGAAACCGAGTCACCTGTGAAAAGATGGGAAATATCTCTCGGAAACCTCAGTTGCTG 221
QY 61 GGAGAGATCAAAAGAACGCCAAGAAAGGAAACCCAGATGAGAAAGAAACCGCAGAA 120
|||||

CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;
 SQ Query Match 28.4%; Score 1068; DB 5; Length 1128;
 Best Local Similarity 98.0%; Pred. No. 1.7e-144;
 Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

QY 1438 CCAAAATTTCAAATGCCAATCTCCCATCCCAATTAATCTTTGAGGTGTAAT 1497
 DB 4 CTACAAATTTAAATGCAACATCTCCCATCCAA-TACCTATTTTCATCTTTGAGGGTAAT 62

QY 1498 CTACTCAATAAATCTGTAGACACAGTACAGACACCTTTGCTTAACCTGACATTTACTTC 1557
 DB 63 CTACTCAATAAATCTGTAGACACAGTACAGACACCTTTGCTTAACCTGACATTTACTTC 122

QY 1558 AATTTTCTTTCTTCTATGTACTGTATATTTTGCATATATAAATTTGCAATAGTCTCAAA 1617
 DB 123 AATTTTCTTTCTTCTATGTACTGTATATTTTGCATATATAAATTTGCAATAGTCTCAAA 182

QY 1618 AATTAATAGTTTTTGACATTTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTACAAAATAA 1677
 DB 183 AATTAATAGTTTTTGACATTTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTACAAAATAA 242

QY 1678 AAAAAAGATGAATGAAGCATATATAATTTGCAATTTTCTCAATTTTCTAGTCAACAGAG 1737
 DB 243 AAAAAAGATGAATGAAGCATATATAATTTGCAATTTTCTCAATTTTCTAGTCAACAGAG 302

QY 1738 ATCGAAGGATTTCTGTTCAAAATATTAGTAAATTTGAAATTTGAAATTTGCTGCTTATTTTG 1797
 DB 303 ATCGAAGGATTTCTGTTCAAAATATTAGTAAATTTGAAATTTGAAATTTGCTGCTTATTTTG 362

QY 1798 TTGCAACACACTAGTTAAATTTAATCTGTAGTATTTCTTACCGAAGGTGATGTGT 1857
 DB 363 TTGCAACACACTAGTTAAATTTAATCTGTAGTATTTCTTACCGAAGGTGATGTGT 422

QY 1858 AGTTTCTGGTTTTAAATTTCAAGCAATCTGAAATTAATCCATTAATTTGCTTTCTTT 1917
 DB 423 AGTTTCTGGTTTTAAATTTCAAGCAATCTGAAATTAATCCATTAATTTGCTTTCTTT 462

QY 1918 CCCAAGAAGTTTTTAAATGATATGCCAGCTTCTTAATTTGGAGACAAAGCCTTAATGA 1977
 DB 483 CCCAAGAAGTTTTTAAATGATATGCCAGCTTCTTAATTTGGAGACAAAGCCTTAATGA 542

QY 1978 CAATGCAATTCATTATATATTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG 2037
 DB 543 CAATGCAATTCATTATATATTTTGTATAGTTACAGTATACAGTTGAGTATCCCTTAG 602

QY 2038 ATGAGATGCTTGGACACAGAGTGTGTTTGGATTTTCAATTTTATTTTGGATTTTGAATA 2097
 DB 603 ATGAGATGCTTGGACACAGAGTGTGTTTGGATTTTCAATTTTATTTTGGATTTTGAATA 662

QY 2098 TTTCCATACATATAATGAGAGAGTTGAAAATGGATTTCAAGTCTAATCATATAAATTCAC 2157
 DB 663 TTTCCATACATATAATGAGAGAGTTGAAAATGGATTTCAAGTCTAATCATATAAATTCAC 722

QY 2158 TTATGTTTGTATACACCTTATCTGAATAGCTGAGGTAATTTTATACAAATTTTAA 2217
 DB 723 TTATGTTTGTATACACCTTATCTGAATAGCTGAGGTAATTTTATACAAATTTTAA 782

QY 2218 TAATTTTATGCTGAAACAGAGTTTGGACATTTGGACATTTGGACATTTGGACATTTGGAC 2277
 DB 783 TAATTTTATGCTGAAACAGAGTTTGGACATTTGGACATTTGGACATTTGGACATTTGGAC 842

QY 2278 TATTTCAAGTCAGTGTCAAAAAGTTTCAAGATGTTAAGCTGGTGTATGATGCTTATGCGAG 2337
 DB 843 TATTTCAAGTCAGTGTCAAAAAGTTTCAAGATGTTAAGCTGGTGTATGATGCTTATGCGAG 902

QY 2338 TGATCCGAGTACTTTGGAGCCAGACAGGTGATCTTTGAGCCAGAGGTTTGAAGC 2397
 DB 903 TGATCCGAGTACTTTGGAGCCAGACAGGTGATCTTTGAGCCAGAGGTTTGAAGC 962

QY 2398 CAGACTGCACAAACACAGTACAGTCTGTTTCTCAAAATTAATTAATAATTTAGCAGGTGT 2457
 DB 963 CAGACTGCACAAACACAGTACAGTCTGTTTCTTCAAAATTAATTAATAATTTAGCAGGTGT 1022

QY 2458 GGTGGTGACACACTGTAGTCCAGGTTACTCAGGAGGCTGAGG-TAGTAGGATTTGTTGAG 2516
 DB 1023 GGTGGTGACACACTGTAGTCCAGGTTACTCAGGAGGCTGAGGTTAGTAGGATTTGTTGAG 1082

QY 2517 ACTGGAGGTTGAGGTTGAACTGAGCCAGGATCTTGCCACCACA 2560
 DB 1083 ACTGGAGGTTGAGGTTGAACTGAGCCAGGATCTTGCCACCACA 1126

RESULT 7
 ID ABV20439 standard; cdna; 1128 BP.
 XX AC ABV20439;
 XX DT 13-SEP-2002 (first entry)
 XX Human prostate expression marker cdna 20430.
 DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX Homo sapiens.
 OS WO200160860-A2.
 PN 23-AUG-2001.
 PD 20-FEB-2001; 2001WO-US0005171.
 PF 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 03-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Schlegel R, Endege WO, Monahan JE;
 PI WPI; 2001-662795/76.
 DR Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX Claim 1; Page 3347; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;
 SQ Query Match 28.4%; Score 1068; DB 5; Length 1128;
 Best Local Similarity 98.0%; Pred. No. 1.7e-144;
 Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

QY 1438 CCAAAATTTCAAATGCCAATCTCCCATCCCAATTAATTTAGGTTGTAAT 1497

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;

Query Match 28.4%; Score 1068; DB 5; Length 1128;
Best Local Similarity 98.0%; Pred. No. 1.7e-144;
Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

QY 1438 CCAAAATTTCAATGCGACATCTCCCATCCCAATTACCTATTTCATCTTTGAGGTGAAT 1497
DB 4 CTACAAATTTAAATGCAACATCTGCCCATCCAA-TACCTATTTCATCTTTGAGGTGAAT 62

QY 1498 CTACTCAATATAACTGTGAAGACCACTGACAGACCCCTTTGCTTAACCTGACATTTTACTTC 1557
DB 63 CTACTCAATATAACTGTGAAGACCACTGACAGACCCCTTTGCTTAACCTGACATTTTACTTC 122

QY 1558 AATTTTCTTTTCTATGACTGATATTTTTCATATAAATCTGAGTAATAGTTTCAA 1617
DB 123 AATTTTCTTTTCTATGACTGATATTTTTCATATAAATCTGAGTAATAGTTTCAA 182

QY 1618 AATTAATAGTTTTCGATCTGCTTTTCTGAGAGAGAAATTCAGAGTCTCAAAATAA 1677
DB 183 AATTAATAGTTTTCGATCTGCTTTTCTGAGAGAGAAATTCAGAGTCTCAAAATAA 242

QY 1678 AAAAAAGATGAAATGAAGCATATATAATTTCTCAATTTTCTAGTCAACAGAGA 1737
DB 243 AAAAAAGATGAAATGAAGCATATATAATTTCTCAATTTTCTAGTCAACAGAGA 302

QY 1738 ATCGAAGGATCTGTTCAATATTTAGTAAAAATTTGAAATTAACCTGTGCTTATATTTG 1797
DB 303 ATCGAAGGATCTGTTCAATATTTAGTAAAAATTTGAAATTAACCTGTGCTTATATTTG 362

QY 1798 TTTTGAACACACTAGTTTAACTTAACTGCTGACTAGTTTCTCTACCGAAGGTGGATGCT 1857
DB 363 TTTTGAACACACTAGTTTAACTTAACTGCTGACTAGTTTCTCTACCGAAGGTGGATGCT 422

QY 1858 AGTTTCTGTTTTTAAAAATTCAGCAAACTGGAAAAATTAATCCATCTAATTTATGCTTTT 1917
DB 423 AGTTTCTGTTTTTAAAAATTCAGCAAACTGGAAAAATTAATCCATCTAATTTATGCTTTT 482

QY 1918 CCAAGAGATTTTAAATGATATGCGAGCTTCTTAATTTGGAGACAAAAGCCTTAATTGA 1977
DB 483 CCAAGAGATTTTAAATGATATGCGAGCTTCTTAATTTGGAGACAAAAGCCTTAATTGA 542

QY 1978 CAATGATTCATATATATTTTCTAGTATACAGATACAGATGAGTATCCCTTAG 2037
DB 543 CAATGATTCATATATATTTTCTAGTATACAGATACAGATGAGTATCCCTTAG 602

QY 2038 ATGAGATGCTTGGGACCAAGGTGTTTGGATTTTCAGATTTATTTTGGATTTTGAATA 2097
DB 603 ATGAGATGCTTGGGACCAAGGTGTTTGGATTTTCAGATTTATTTTGGATTTTGAATA 662

QY 2098 TTTCCATATATATAGAGAGAGTTGGAAAAATGGGATTCAGTCTAATCAATAAATTCAC 2157
DB 663 TTTCCATATATATAGAGAGAGTTGGAAAAATGGGATTCAGTCTAATCAATAAATTCAC 722

QY 2158 TTATGTTTTCATATACACCTTATCTGAATAGCCTGAAGGTAATTTTATACATAATTTTAA 2217
DB 723 TTATGTTTTCATATACACCTTATCTGAATAGCCTGAAGGTAATTTTATACATAATTTTAA 782

QY 2218 TAATTTATGCTTGAACACAGTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 2277
DB 783 TAATTTATGCTTGAACACAGTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 842

DB 783 TAATTTATGCTTGAACACAGTTTCGCAATTTGGACCATCAGAAAGCAGAGTGTAC 842

QY 2278 TATTTCAAGTCAGTCTCAAAAAGTTTCAGATGTTTAAAGTGTGATGCAAGTTTCATGCCAG 2337
DB 843 TATTTCAAGTCAGTCTCAAAAAGTTTCAGATGTTTAAAGTGTGATGCAAGTTTCATGCCAG 902

QY 2338 TGATCCGAGTACTTTGGGAAGCCAAAGACAGAGTGGATCTCTTGAGCCCGAGAGTTTGAGGC 2397
DB 903 TGATCCGAGTACTTTGGGAAGCCAAAGACAGAGTGGATCTCTTGAGCCCGAGAGTTTGAGGC 962

QY 2398 CAGATGTCACACACAGTGAACCTGCTTTCTACAAAATAATTTAAAAAATTCAGAGTGT 2457
DB 963 CAGACTGTCACACACAGTGAACCTGCTTTCTCAATAATAATTTAAAAAATTCAGCAGGTGT 1022

QY 2458 GGTGTCACACCTCTAGTCTCCAGTACTCAGGAGCTGAGG-TAGTAGGATTTGTTGAG 2516
DB 1023 GGTGTCACACCTCTAGTCTCCAGTACTCAGGAGCTGAGGATTTGAGATTTGTTGAG 1082

QY 2517 ACTGGGAGTTTGAGGCTGAACTGAGCCAGGATCTTTGCCACCCACA 2560
DB 1083 ACTGGGAGTTTGAGGCTGAACTGAGCCAGGATCTTTGCCACCCACA 1126

RESULT 6
ABV20806
ID ABV20806 standard; cDNA; 1128 BP.
XX AC ABV20806;
XX DT 13-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 20797.
XX DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW Pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US0051171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JB;
XX PI WPI; 2001-662795/76.
XX DR Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 3412-3413; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)

PI	XX	Schlegel R, Endege WO, Monahan JE;	2098	TTTCCATACATATAATGAGAGATTGGAAATGGGATTCAGTCTAATCATATAAATTCAC	2151
XX	XX	WPI; 2001-662795/76.	663	TTTCCATACATATAATGAGAGATTGGTAATGGGATTCAGTCTAATCATATAATTCAC	722
XX	XX	Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.	2158	TTATGTTTGATATACACCTTATCTGAATAGCCTGAAAGTAAATTTATACAAATTTTAAA	2217
PT	PT	Claim 1; Page 5346-5347; 11750pp; English.	723	TTATGTTTGATATACACCTTATCTGAATAGCCTGAAAGTAAATTTATACAAATTTTAAA	782
XX	XX	The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV0010-ABV6213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker	2218	TAATTTTATGCTGAAACAGAGATTGGGCACATTTGGACCATCAGAAAGCAGAGTGTCC	2277
PS	PS	Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;	783	TAATTTTATGCTGAAACAGAGATTGGGCACATTTGGACCATCAGAAAGCAGAGTGTCC	842
CC	CC	Query Match 28.4%; Score 1068; DB 5; Length 1128;	2278	TATTTTCAAGTCAGTCTCTCAAAAAGTTTCAGATGTTAAGCTGGTATGATGCGAGTTC	2337
CC	CC	Best Local Similarity 98.0%; Pred. No. 1.7e-144;	843	TATTTTCAAGTCAGTCTCTCAAAAAGTTTCAGATGTTAAGCTGGTATGATGCGAGTTC	902
CC	CC	Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;	2338	TGATCCGAGTACTTTGGGAAGCCAGACAGTGGATCTTTTGAGCCAGAGTTTGAGGC	2397
CC	CC	1438 CCAAAATTCAAATGCCAACATCTCCCATCCGAATTAACCTATTTCATCTTGAGTGAAT	903	TGATCCGAGTACTTTGGGAAGCCAGACAGTGGATCTTTTGAGCCAGAGTTTGAGGC	962
CC	CC	4 CTACAAATTTAAATGCAACATCTGCCATCCAA-TACCTATTTCATCTTTGAGGGTAAAT	2398	CAGACTGTCACACACAGTGCAGACCTCGTTTCTACAAATTAATTAATAAATTAGCCAGTGT	2457
CC	CC	1498 CTACTCAATAAACTGTGAAGACCACTGACAGACCTTTGCTAACCTGCATTTACTTC	963	CAGACTGTCACACACAGTGCAGACCTCGTTTCTTCAATAAATTAATAAATTAGCCAGTGT	1022
CC	CC	63 CTACTCAATAAACTGTGAAGACCACTGACAGACCTTTGCTAACCTGCATTTACTTC	2458	GGTGTGTCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGG-TAGTAGGATTTGTTGAG	2516
CC	CC	1558 AATTTTCTTTTCTATGTAATGTAATTTTGGATATATAAACTTGCAGTAAATAGTCAAA	1023	GGTGTGTCACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGATTTGGATTTGTTGAG	1082
CC	CC	123 AATTTTCTTTTCTATGTAATGTAATTTTGGATATATAAACTTGCAGTAAATAGTCAAA	2517	ACTGGAGGTTGAGGCTGAACTGAGCCAGAGTCTTTGCGACCA	2560
CC	CC	1618 AATTATAGTTTTTGACATTCGCTTTCTGAGAGAGAAATGAAAGTGCACAAATAA	1083	ACTGGAGGTTGAGGCTGAACTGAGCCAGAGTCTTTGCGACCA	1126
CC	CC	183 AATTATAGTTTTTGACATTCGCTTTCTGAGAGAGAAATGAAAGTGCACAAATAA	RESULT 5		
CC	CC	1678 AAAAGATCAAAATGAAGCATATATAATTTGTCATTTTCAATTTTCTAGTCACACAGA	ABV20617		
CC	CC	243 AAAAGATCAAAATGAAGCATATATAATTTGTCATTTTCAATTTTCTAGTCACACAGA	ID	ABV20617 standard; cDNA; 1128 BP.	
CC	CC	1738 ATGGAAGATTCGTTTCAAAATATAGTAAAAATGAAATTAACCTTGTGCTTATATTTTG	XX	ABV20617;	
CC	CC	303 ATGGAAGATTCGTTTCAAAATATAGTAAAAATGAAATTAACCTTGTGCTTATATTTTG	XX	AC	
CC	CC	1798 TTGGAACACATAGTAAATTTAACTGTGACTAGTATCTCTACCGAAGTGAATGTGT	XX	ABV20617;	
CC	CC	363 TTGGAACACATAGTAAATTTAACTGTGACTAGTATCTCTACCGAAGTGAATGTGT	DT	13-SEP-2002 (first entry)	
CC	CC	1858 AGTTTCTGTTTTTAAATTCAGCAAACTGGAATAATCCATCTAATTTGCTTTCTTT	XX	Human prostate expression marker cDNA 20608.	
CC	CC	423 AGTTTCTGTTTTTAAATTCAGCAAACTGGAATAATCCATCTAATTTGCTTTCTTT	DE	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;	
CC	CC	1918 CCAAGAAATTTTTTAATGATGCGAGCTTCTTAATTTGGAGCAAAAGCCTTAATGA	XX	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;	
CC	CC	483 CCAAGAAATTTTTTAATGATGCGAGCTTCTTAATTTGGAGCAAAAGCCTTAATGA	XX	Homo sapiens.	
CC	CC	1978 CAATGCAATTCATATATATTTTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG	XX	WO200160860-A2.	
CC	CC	543 CAATGCAATTCATATATATTTTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG	PD	23-AUG-2001.	
CC	CC	2038 ATGAGATGCTGGGACCAAGTGTGTTGGATTTTCAAGTATTTATTTTGGATTTGGAATA	XX	20-FEB-2001; 2001WO-US005171.	
CC	CC	603 ATGAGATGCTGGGACCAAGTGTGTTGGATTTTCAAGTATTTATTTTGGATTTTGGAAATA	XX	17-FEB-2000; 2000US-0183319P.	
CC	CC		XX	16-MAR-2000; 2000US-0189862P.	
CC	CC		XX	25-MAY-2000; 2000US-0207454P.	
CC	CC		XX	09-JUN-2000; 2000US-0211314P.	
CC	CC		XX	18-JUL-2000; 2000US-0219007P.	
CC	CC		XX	13-DEC-2000; 2000US-0255281P.	
CC	CC		XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.	
CC	CC		XX	Schlegel R, Endege WO, Monahan JE;	
CC	CC		XX	WPI; 2001-662795/76.	
CC	CC		XX	Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.	
CC	CC		XX	Claim 1; Page 3379; 11750pp; English.	

XX PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Schlegel R, Endege WO, Monahan JE;
 PI WPI; 2001-662795/76.
 DR Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX Claim 1; Page 5285-5286; 11750pp; English.
 PS The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;
 Query Match 28.4%; Score 1068; DB 5; Length 1128;
 Best Local Similarity 98.0%; Pred. No. 1.7e-144;
 Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;
 QY 1438 CCAAAATTTCAAAATGCCAATCTCCCATCCCAATACCTATTTTCATCTTTGAGGTGAAT 1497
 DB 4 CTACAAATTTAAATGCAACATCTGCCCATCCAA-TACCTATTTTCATCTTTGAGGTGAAT 62
 QY 1498 CTACTCAATAAAGTGTGAAGACCTGACGACCCCTTTGCTAACTGACATTTACTTC 1557
 DB 63 CTACTCAATAAAGTGTGAAGACCTGACGACCCCTTTGCTAACTGACATTTACTTC 122
 QY 1558 AATTTTCTTTCTATGTACTGGATATTTTGTGATATAAATCTGAGTAATAGTTTCAAA 1617
 DB 123 AATTTTCTTTCTATGTACTGGATATTTTGTGATATAAATCTGAGTAATAGTTTCAAA 182
 QY 1618 AATTAATAGTTTGTGACATGCTTTCTGAGAGAGAAATGAAAGTGTCAAAATATAA 1677
 DB 183 AATTAATAGTTTGTGACATGCTTTCTGAGAGAGAAATGAAAGTGTCAAAATATAA 242
 QY 1678 AAAAAAGTCAATGAGCAGATATAATCTGCAATTTTCTAGTCAACAGAGA 1737
 DB 243 AAAAAAGTCAATGAGCAGATATAATCTGCAATTTTCTAGTCAACAGAGA 302
 QY 1738 ATCGAAGGATCTGTTCAATATTTAGTAAAAATGAAAAATGAACTTGTGCTTATTTTG 1797
 DB 303 ATCGAAGGATCTGTTCAATATTTAGTAAAAATGAAAAATGAACTTGTGCTTATTTTG 362
 QY 1798 TTGCAACACACTAGTTAATTTAACTGTGACTAGTTATCTCTACCAAGGTGATGTGT 1857
 DB 363 TTGCAACACACTAGTTAATTTAACTGTGACTAGTTATCTCTACCAAGGTGATGTGT 422
 QY 1858 AGTTTCTGTTTAAAAATTCAGCAAACTGGAATAAATCCATCTAATATGCTTTCTTT 1917
 DB 423 AGTTTCTGTTTAAAAATTCAGCAAACTGGAATAAATCCATCTAATATGCTTTCTTT 482
 QY 1918 CCAAGAAGTTTTTAAATGATGCGAGCTTCTTAATTTGGAGACAAAAGCCCTTAATGA 1977

DB 483 CCCAAGAAAGTTTTTAAATGATGCGAGCTTCTTAATTTGGAGACAAAAGCCCTTAATGA 542
 QY 1978 CAATGCATTTCATATATATTTTTTTTGTATAGTTTACAGTATACAGTTGAGTATCCCTTAG 2037
 DB 543 CAATGCATTTCATATATATTTTTTTTGTATAGTTTACAGTATACAGTTGAGTATCCCTTAG 602
 QY 2038 ATGAGATGCTTGGGACCAAGAGTGTTTTGGATTTTTCAGATTTTATTTTGGATTTTGAATA 2097
 DB 603 ATGAGATGCTTGGGACCAAGAGTGTTTTGGATTTTTCAGATTTTATTTTGGATTTTGAATA 662
 QY 2098 TTTCCATACATATAATGAGAGAGTTGGAAAAATGGGATTTCAAGTCTAATCATATAAATTCAC 2157
 DB 663 TTTCCATACATATAATGAGAGAGTTGGTAAATGGGATTTCAAGTCTAATCATATAAATTCAC 722
 QY 2158 TTATGTTTGCATATACACCTTATCTGAATAGCTGAGGTAATTTTATACAAATATTTTAA 2217
 DB 723 TTATGTTTGCATATACACCTTATCTGAATAGCTGAGGTAATTTTATACAAATATTTTAA 782
 QY 2218 TAAATTTTATGCTGAAACAGAGTTTGGCCACATTTGGCCACATCAGAAAGCAGAGTGTCTAC 2277
 DB 783 TAAATTTTATGCTGAAACAGAGTTTGGCCACATTTGGCCACATCAGAAAGCAGAGTGTCTAC 842
 QY 2278 TATTTCAAGTCAAGTCTCAAAAAGTTTCAAGTCTTAAAGTGTGATGCGAGTTTCATGCCAG 2337
 DB 843 TATTTCAAGTCAAGTCTCAAAAAGTTTCAAGTCTTAAAGTGTGATGCGAGTTTCATGCCAG 902
 QY 2338 TGATCCGAGTACTTTGGGAAGCCAGACAGAGTGGATCTCTTGAGCCCGAGGAGTTTGAGGC 2397
 DB 903 TGATCCGAGTACTTTGGGAAGCCAGACAGAGTGGATCTCTTGAGCCCGAGGAGTTTGAGGC 962
 QY 2398 CAGACTGCAACACACAGTGAGACCTCGTTTCTACAAATTAATTAATAAATTTAGCCAGGTGT 2457
 DB 963 CAGACTGCAACACACAGTGAGACCTCGTTTCTCAAAATTAATTAATAAATTTAGCCAGGTGT 1022
 QY 2458 GGTGTGTCACACTGTAGTCCAGGACTCAGAGGCTGAGGCTAGGCTAGGCTAGGCTAGGCTAGG 2516
 DB 1023 GGTGTGTCACACTGTAGTCCAGGACTCAGAGGCTGAGGCTAGGCTAGGCTAGGCTAGGCTAGG 1082
 QY 2517 ACTGGGAGGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2560
 DB 1083 ACTGGGAGGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1126
 XX RESULT 4
 XX ABV26458
 XX ID ABV26458 standard; cDNA; 1128 BP.
 XX AC ABV26458;
 XX DT 16-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 26449.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX XX WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US005171.
 XX PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.

Homo sapiens.

WO200160860-A2.

23-AUG-2001.

20-FEB-2001; 2001WO-US005171.

17-FEB-2000; 2000US-0183319P.

16-MAR-2000; 2000US-0189862P.

25-MAY-2000; 2000US-0207454P.

09-JUN-2000; 2000US-0211314P.

18-JUL-2000; 2000US-0219007P.

13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

Claim 1; Page 3320-3321; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABY00010-ABY62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;

Query Match 28.4%; Score 1068; DB 5; Length 1128;

Best Local Similarity 98.0%; Pred. No. 1.7e-144;

Matches 1102; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

QY	1438	CCAAATTTCAATGCCAACATCTCCCATCCCAATTACCTATTTTCATCTTTGAGGTGAAT	1497
DB	4	CTACAAATTTAATGACATCTGCCATCCAA-TACCTATTTCACTTTTGAGGGTAAT	62
QY	1498	CTACTCAATAAAGTGTGAAGACAGTGCAGACCCCTTTGCTAATGCTGATTTACTTC	1557
DB	63	CTACTCAATAAAGTGTGAAGACAGTGCAGACCCCTTTGCTAATGCTGATTTACTTC	122
QY	1558	AATTTTCTTTTCTATGCTGATATTTTTCATATAAATTTGAGTAATGCTCAAA	1617
DB	123	AATTTTCTTTTCTATGCTGATATTTTTCATATAAATTTGAGTAATGCTCAAA	182
QY	1618	AATTAATAGTTTTTGACATGGCTTTTCTGAGAGAGAAATTCAGAGTGCACAAATAA	1677
DB	183	AATTAATAGTTTTTGACATGGCTTTTCTGAGAGAGAAATTCAGAGTGCACAAATAA	242
QY	1678	AAAAAGTAAGTAAGCATATAATTTGCAATTTTTCATTTTCTAGTCAACAGAGA	1737
DB	243	AAAAAGTAAGTAAGCATATAATTTGCAATTTTTCATTTTCTAGTCAACAGAGA	302
QY	1738	ATCGAAGGATTTCTGCAATATATAGTAAATTTGAAATTAATTTGCTGCTTATTTTG	1797
DB	303	ATCGAAGGATTTCTGCAATATATAGTAAATTTGAAATTAATTTGCTGCTTATTTTG	362

QY	1798	TTTGCAACACACTAGTTAAATTTAACTGTGCTAGTTATCTCTACGAAGGTGATGTGT	1857
DB	363	TTTGCAACACACTAGTTAAATTTAACTGTGCTAGTTATCTCTACGAAGGTGATGTGT	422
QY	1858	AGTTTCTGGTTTAAATTTCAAGCAACTGCAAAATTAATCCATCTAATTAATGCTTTCT	1917
DB	423	AGTTTCTGGTTTAAATTTCAAGCAACTGCAAAATTAATCCATCTAATTAATGCTTTCT	482
QY	1918	CCCAAGAAAGTTTAAATGATATGCCAGCTTCTTAATTTGGAGACAAAGCCTTAATTTGA	1977
DB	483	CCCAAGAAAGTTTAAATGATATGCCAGCTTCTTAATTTGGAGACAAAGCCTTAATTTGA	542
QY	1978	CAATGCATTCATTATATATTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG	2037
DB	543	CAATGCATTCATTATATATTTTGTATAGTTACAGTATACGAGTTGAGTATCCCTTAG	602
QY	2038	ATCAGATGCTTTGGGACCGAAGTGTTCGATTTTCAGATTTTGGATTTTGGATTTGGAATA	2097
DB	603	ATCAGATGCTTTGGGACCGAAGTGTTCGATTTTCAGATTTTGGATTTTGGATTTGGAATA	662
QY	2098	TTTCCATACATATAATGAGAGAGTTTGGAAAATGGGATTTCAAGTCTAATCATATAAATTCAC	2157
DB	663	TTTCCATACATATAATGAGAGAGTTTGGAAAATGGGATTTCAAGTCTAATCATATAAATTCAC	722
QY	2158	TTATGTTTGTATATACACCTTATCTGAATAGCTTGAAGGTAATTTTATACAAATTTTAA	2217
DB	723	TTATGTTTGTATATACACCTTATCTGAATAGCTTGAAGGTAATTTTATACAAATTTTAA	782
QY	2218	TAATTTTATGCTTGAACAGAGTTTGGGACATTTGGCCATTCAGACCATCAGAAAGCAGAGTGTAC	2277
DB	783	TAATTTTATGCTTGAACAGAGTTTGGGACATTTGGCCATTCAGACCATCAGAAAGCAGAGTGTAC	842
QY	2278	TATTTCAAGTCACTGCTCAAAAAGTTTTCAGATGTTAAGCTGGTGTGATGCTCAGTTTCATGCCAG	2337
DB	843	TATTTCAAGTCACTGCTCAAAAAGTTTTCAGATGTTAAGCTGGTGTGATGCTCAGTTTCATGCCAG	902
QY	2338	TGATCCGAGTACTTTGGGAAGCCAGACAGGTGGATCTCTTGAGCCAGAGTTTGAAGGC	2397
DB	903	TGATCCGAGTACTTTGGGAAGCCAGACAGGTGGATCTCTTGAGCCAGAGTTTGAAGGC	962
QY	2398	CAGACTGCACACACAGTGCAGCTCTGTTTCTACAAATTAATTAATAAATTTAGCCAGGTGT	2457
DB	963	CAGACTGCACACACAGTGCAGCTCTGTTTCTACAAATTAATTAATAAATTTAGCCAGGTGT	1022
QY	2458	GGTGTGTCACACCTGTAGTCCAGGCTACTCAGAGGCTGAGG-TAGTAGGATTTGTTTGG	2516
DB	1023	GGTGTGTCACACCTGTAGTCCAGGCTACTCAGAGGCTGAGGTTATTGGGATTTGTTTGG	1082
QY	2517	ACTGGGAGGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT	2560
DB	1083	ACTGGGAGGTTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT	1126

RESULT 3

ABV26130

ID ABV26130 standard; cDNA; 1128 BP.

XX ABV26130;

AC ABV26130;

XX 16-SEP-2002 (first entry)

DT 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 26121.

DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX Homo sapiens.

XX WO200160860-A2.

PN 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX

PF

Db 1741 GAAGGATCTGTTCAAAATATTAGTAAATAATGAAATAATTAACCTGTGCTTATATTGTTT 1800
 QY 1801 GCAACACACTAGTTAAATTAACCTGTGACTAGTTATCTCTACCGAAGGTGGATGTAGT 1860
 Db 1801 GCAACACACTAGTTAAATTAACCTGTGACTAGTTATCTCTACCGAAGGTGGATGTAGT 1860
 QY 1861 TTCTGGTTTAAATTAACCAAGCAAACTGGAAATAATTAATCCATCTAATTAATGCTTCTCTTCCC 1920
 Db 1861 TTCTGGTTTAAATTAACCAAGCAAACTGGAAATAATTAATCCATCTAATTAATGCTTCTCTTCCC 1920
 QY 1921 AAGGAGTTTAAATTAATGATATGCGAGCTTCTTAATTTGGAGACAAAGGCTTAAATTTGACAA 1980
 Db 1921 AAGGAGTTTAAATTAATGATATGCGAGCTTCTTAATTTGGAGACAAAGGCTTAAATTTGACAA 1980
 QY 1981 TGCATTCAATATATATTTTTTGTATAGTTTACAGTATACAGTATGAGTATCCCTTTAGATG 2040
 Db 1981 TGCATTCAATATATATTTTTTGTATAGTTTACAGTATACAGTATGAGTATCCCTTTAGATG 2040
 QY 2041 AGATGCTTGGACCAAGAGTGTGTTGGATTTTCAGATTTTATTTTGGATTTTGGATATTT 2100
 Db 2041 AGATGCTTGGACCAAGAGTGTGTTGGATTTTCAGATTTTATTTTGGATTTTGGATATTT 2100
 QY 2101 CCATACATATATGAGAGTGTGGAATGGATTTCAAGTCTTAATCATATAAATTCACCTTA 2160
 Db 2101 CCATACATATATGAGAGTGTGGAATGGATTTCAAGTCTTAATCATATAAATTCACCTTA 2160
 QY 2161 TGTGTTGATATACACCTTATCTGAATAGCTGAAGGTAAATTTTATACAAATTTTAAATAA 2220
 Db 2161 TGTGTTGATATACACCTTATCTGAATAGCTGAAGGTAAATTTTATACAAATTTTAAATAA 2220
 QY 2221 TTTTATGCTGAAACAGAGTTTGGCGACATTTGGACATCAGAAAGCAAGTGTCACTAT 2280
 Db 2221 TTTTATGCTGAAACAGAGTTTGGCGACATTTGGACATCAGAAAGCAAGTGTCACTAT 2280
 QY 2281 TTCAAGTCACTGCTCAAAAAGTTTCAGATGTTTAAAGTGTGATGCAAGTTCATGCGAGTGA 2340
 Db 2281 TTCAAGTCACTGCTCAAAAAGTTTCAGATGTTTAAAGTGTGATGCAAGTTCATGCGAGTGA 2340
 QY 2341 TCCGAGTACTTTGGGAAGCCAAAGACAGGTGGATCTCTTGAGCCCGAGGATTTTGAGGCCAG 2400
 Db 2341 TCCGAGTACTTTGGGAAGCCAAAGACAGGTGGATCTCTTGAGCCCGAGGATTTTGAGGCCAG 2400
 QY 2401 ACTGCACAAACAGTGGAGCTGTTTCTACAAATAATTAATAATTTAGCCAGGTGTGTT 2460
 Db 2401 ACTGCACAAACAGTGGAGCTGTTTCTACAAATAATTAATAATTTAGCCAGGTGTGTT 2460
 QY 2461 GGTGCACACCTGTAGTCCAGGTACTCAGAGGCTGAGGTAGTATGTTTGGAGACTG 2520
 Db 2461 GGTGCACACCTGTAGTCCAGGTACTCAGAGGCTGAGGTAGTATGTTTGGAGACTG 2520
 QY 2521 GGAGGTGAGGCTGAACTGAGCCAGGATCTTGCCACCAATTCAGCTTTGGGCAACAGAG 2580
 Db 2521 GGAGGTGAGGCTGAACTGAGCCAGGATCTTGCCACCAATTCAGCTTTGGGCAACAGAG 2580
 QY 2581 TGAGACCTGCTCAAAAAGTTTAAAGTGGAGGGAACCAATAGAAATTTTGCATCTTCAG 2640
 Db 2581 TGAGACCTGCTCAAAAAGTTTAAAGTGGAGGGAACCAATAGAAATTTTGCATCTTCAG 2640
 QY 2641 ATTAGGGATTTCAACCTGTACTGACCTTTTATGATCAATGACAGCAATTAATCAATAGGTTG 2700
 Db 2641 ATTAGGGATTTCAACCTGTACTGACCTTTTATGATCAATGACAGCAATTAATCAATAGGTTG 2700
 QY 2701 GACTCCAGATAAATCAATTTGCTGTATACAAATTTTGCCTCTCTAATTCACAGCAATCTTAT 2760
 Db 2701 GACTCCAGATAAATCAATTTGCTGTATACAAATTTTGCCTCTCTAATTCACAGCAATCTTAT 2760
 QY 2761 GGCCTCTTGGGTGATTTAAAGTGGAGGGAACCAATAGAAATTTTGCATCTTCAG 2820
 Db 2761 GGCCTCTTGGGTGATTTAAAGTGGAGGGAACCAATAGAAATTTTGCATCTTCAG 2820
 QY 2821 AAAGTCATTTCTGTCAAAATATGTCAAGTCTGTAGATATTAGCCAAATTTTAGGAAATATGAC 2880

Db 2821 AAAGTCATTTCTGTCAAAATATGTCAAGTCTGTAGATATTAGCCAAATTTTAGGAAATATGAC 2880
 QY 2881 AAAGTCATTTCTGTCAAAATATGTCAAGTCTGTAGATATTAGCCAAATTTTAGGAAATATGAC 2880
 Db 2881 AAAGTCATTTCTGTCAAAATATGTCAAGTCTGTAGATATTAGCCAAATTTTAGGAAATATGAC 2880
 QY 2941 AATAAAAATTAATTTTAAATTTTCAGTAAACAATCTGGAATTTATCAGAGAAGGGCAACAATA 3000
 Db 2941 AATAAAAATTAATTTTAAATTTTCAGTAAACAATCTGGAATTTATCAGAGAAGGGCAACAATA 3000
 QY 3001 GGTAAATAAAACAGTATTTGATTTGAGAGAAACGTTGAAATCCAGAGACATCAATGTCTT 3060
 Db 3001 GGTAAATAAAACAGTATTTGATTTGAGAGAAACGTTGAAATCCAGAGACATCAATGTCTT 3060
 QY 3061 CTGGTGGTTTCAACATTAAGCCACAGAGATGCTTAATCTTTTCCGAGATCTAGTTTTCAG 3120
 Db 3061 CTGGTGGTTTCAACATTAAGCCACAGAGATGCTTAATCTTTTCCGAGATCTAGTTTTCAG 3120
 QY 3121 CAAAGCAGGATTTAAGAAATGTAATCTTATGTTGTTTATGAAGAAACAATAGAAATCAT 3180
 Db 3121 CAAAGCAGGATTTAAGAAATGTAATCTTATGTTGTTTATGAAGAAACAATAGAAATCAT 3180
 QY 3181 GCTGTATAGTGTCTTTTAACTGTAAATTTTGTGAAGCTTATCTTTTATGCAATATAA 3240
 Db 3181 GCTGTATAGTGTCTTTTAACTGTAAATTTTGTGAAGCTTATCTTTTATGCAATATAA 3240
 QY 3241 ATTTGAAACATTTTACATTTGTTTATATTTTAACTGATTTTACTCAAGTGTGATTTATATAC 3300
 Db 3241 ATTTGAAACATTTTACATTTGTTTATATTTTAACTGATTTTACTCAAGTGTGATTTATATAC 3300
 QY 3301 AAGAAATGTAAACCACTGTAAGGGTAGAGTTTATAAGAAATTTTGTCAAATGTATTCACCCA 3360
 Db 3301 AAGAAATGTAAACCACTGTAAGGGTAGAGTTTATAAGAAATTTTGTCAAATGTATTCACCCA 3360
 QY 3361 TGTAGTCACTCTCTTATGAAGAGACAGAACAGTACATCTCCAGAGAGTTCCACAGTG 3420
 Db 3361 TGTAGTCACTCTCTTATGAAGAGACAGAACAGTACATCTCCAGAGAGTTCCACAGTG 3420
 QY 3421 CTCCTTTTCCCTGAGTTTCCAGGTCCTGGCAACCAATGATCTGCTTCGTATAATTTATAA 3480
 Db 3421 CTCCTTTTCCCTGAGTTTCCAGGTCCTGGCAACCAATGATCTGCTTCGTATAATTTATAA 3480
 QY 3481 CTGTTCTAGATATTTGTAGCAATGTACCTTTTCCATATTTTATTTTGTGTGTGTAAGGCTT 3540
 Db 3481 CTGTTCTAGATATTTGTAGCAATGTACCTTTTCCATATTTTATTTTGTGTGTGTAAGGCTT 3540
 QY 3541 CTTTGTAGTCAATTAATAATTTTGTAGATCTATGTTTAAATGTTTCTATCAGTAGTTGT 3600
 Db 3541 CTTTGTAGTCAATTAATAATTTTGTAGATCTATGTTTAAATGTTTCTATCAGTAGTTGT 3600
 QY 3601 ACATCTTACTGCTCAGCATATCACCATATAGATATCTATAATTTGTTAACTTAATCA 3660
 Db 3601 ACATCTTACTGCTCAGCATATCACCATATAGATATCTATAATTTGTTAACTTAATCA 3660
 QY 3661 CTGATGGATATGAGGATATTTTAAAGTTTGTGACATTTGAATAAAGTGGCTATAAATGAA 3720
 Db 3661 CTGATGGATATGAGGATATTTTAAAGTTTGTGACATTTGAATAAAGTGGCTATAAATGAA 3720
 QY 3721 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3762
 Db 3721 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3762

RESULT 2
 ABV20298

ID ABV20298 standard; cDNA; 1128 BP.

XX
 AC ABV20298;

XX
 DT 13-SEP-2002 (first entry)

XX
 DE Human prostate expression marker cDNA 20289.

XX

XX The invention relates to isolated nucleic acid (ABN81319-ABN81324),
 CC corresponding to genes differentially expressed in mast cells following
 CC activation or in patients with allergic hypersensitivity disease, (I)
 CC that encodes proteins (ABB7569-ABB7575) (II) or a protein fragment of
 CC (II) if at least 6 amino acids, (II) is useful for identifying binding
 CC partners. (I) or (II) is useful for diagnosing or treating a disease
 CC state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
 CC urticaria or atopic dermatitis or mastocytosis) in a subject which
 CC involves determining the level of expression of (I) or (II). A computer
 CC system, comprising a database containing information identifying the
 CC expression level in a tissue or at least one mast cell of (I), is useful
 CC for presenting information to identify the relative expression level of
 CC (I). (II) is used as a marker to detect, diagnose or identify an allergic
 CC response in a patient. The protein can also serve as a target that
 CC modulate gene expression or activity and as an antigen to raise
 CC polyclonal or monoclonal antibodies. (II) is useful for identifying
 CC agents that modulate expression of the protein or agents, such as
 CC agonists or antagonists. The agonists or antagonists are useful for
 CC modulating biological activity and function of (II) and thus are useful
 CC for alleviating disease conditions such as allergic hypersensitivity,
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis
 XX
 SQ Sequence 3762 BP; 1220 A; 672 C; 680 G; 1190 T; 0 U; 0 Other;
 Query Match 100.0%; Score 3762; DB 6; Length 3762;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGAAACCGAGTCACTGTGAAAGATGGAAATATCTCTCGGAAACCTCAGTTGGCTG 60
 DB 1 GAGAAACCGAGTCACTGTGAAAGATGGAAATATCTCTCGGAAACCTCAGTTGGCTG 60
 QY 61 GAGAGAGATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGGAAAGAAACCGCAGGAA 120
 DB 61 GAGAGAGATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGGAAAGAAACCGCAGGAA 120
 QY 121 ATGACTACATTTGAAAGAAACCTTCAAGATCAAGATGAAGAAAGCCAGAGTTTCATCC 180
 DB 121 ATGACTACATTTGAAAGAAACCTTCAAGATCAAGATGAAGAAAGCCAGAGTTTCATCC 180
 QY 181 ACTTCTAATCAGGAAACCGAGAAATGGCAGTGGTTCGAAAGAGTGTCTACACTGTCAAT 240
 DB 181 ACTTCTAATCAGGAAACCGAGAAATGGCAGTGGTTCGAAAGAGTGTCTACACTGTCAAT 240
 QY 241 AATCAGATCCCCATCAGAGATCCTCCTGAGTCCAAATGATGATGGCTATGAGAACAT 300
 DB 241 AATCAGATCCCCATCAGAGATCCTCCTGAGTCCAAATGATGATGGCTATGAGAACAT 300
 QY 301 GACTCCCTCAAGAGAAAGTGAGACAGTTTATAGAGAAAGTTCAGAGACAGAAATATGCCCTT 360
 DB 301 GACTCCCTCAAGAGAAAGTGAGACAGTTTATAGAGAAAGTTCAGAGACAGAAATATGCCCTT 360
 QY 361 CTATAGACATCTGTAGTAGGCTTGTCTCTGACCCCATCAGCATGATTAAGTTGTG 420
 DB 361 CTATAGACATCTGTAGTAGGCTTGTCTCTGACCCATCAGCATGATTAAGTTGTG 420
 QY 421 TTTCCACACTAAATCCTCAAGCTGCTTTATACCTTCCAGCAATGAAGCAATGAGAA 480
 DB 421 TTTCCACACTAAATCCTCAAGCTGCTTTATACCTTCCAGCAATGAAGCAATGAGAA 480
 QY 481 TAGCAGACTCTGGCGAAGTTGTTCAACCTGAGCAGTGCATGAACATTTCTCTGGCTA 540
 DB 481 TAGCAGACTCTGGCGAAGTTGTTCAACCTGAGCAGTGCATGAACATTTCTCTGGCTA 540
 QY 541 AAGTTTGAAGAAATATATCTTATATATATCTCTAGGCAACTCTGATATGGCATCTCTG 600
 DB 541 AAGTTTGAAGAAATATATCTTATATATATCTCTAGGCAACTCTGATATGGCATCTCTG 600
 QY 601 TGGCTTAGGTAATCATAGAAATTCACAAATGACCTAAATATTTCTATGTTTTTTCG 660
 DB 601 TGGCTTAGGTAATCATAGAAATTCACAAATGACCTAAATATTTCTATGTTTTTTCG 660

QY 661 TTGTAAAGTTTGAGGACATGAGAGTGTATAAAAAAACTTTCTTAGGCAATAATATGTAAAA 720
 DB 661 TTGTAAAGTTTGAGGACATGAGAGTGTATAAAAAAACTTTCTTAGGCAATAATATGTAAAA 720
 QY 721 TGAATAAATATTTCTAATCCCTGACTACTGAATGAGACCTCTTCTAGGCCAAAGAGA 780
 DB 721 TGAATAAATATTTCTAATCCCTGACTACTGAATGAGACCTCTTCTAGGCCAAAGAGA 780
 QY 781 CCTCAGATGAACCTGAAAGACTGAATTCCTGGCCATGATAGAGGGAGGTGAGACACACC 840
 DB 781 CCTCAGATGAACCTGAAAGACTGAATTCCTGGCCATGATAGAGGGAGGTGAGACACACC 840
 QY 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGACCCAGGATGAGTCAAGACT 900
 DB 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGACCCAGGATGAGTCAAGACT 900
 QY 901 GATGAATAACAGTATTTGCGCAATAGAGTCCCAATTCCAACTGACTCTGGTTAGAT 960
 DB 901 GATGAATAACAGTATTTGCGCAATAGAGTCCCAATTCCAACTGACTCTGGTTAGAT 960
 QY 961 CACACACTGTCTGAGGGAATTCATCTATGAGACTTTGTCTACATAACAGAGACCTTGTT 1020
 DB 961 CACACACTGTCTGAGGGAATTCATCTATGAGACTTTGTCTACATAACAGAGACCTTGTT 1020
 QY 1021 TCCACAAACCCCTTATTTAGCTTAAAGCAATCTTTCTACTGACTTCTTAAGTCTTTTGA 1080
 DB 1021 TCCACAAACCCCTTATTTAGCTTAAAGCAATCTTTCTACTGACTTCTTAAGTCTTTTGA 1080
 QY 1081 CAAAGCTTAACTCTTCAACCAATTCGCAATAGAGTCCCAATTCCAACTGACTCTGGTTAGAT 1140
 DB 1081 CAAAGCTTAACTCTTCAACCAATTCGCAATAGAGTCCCAATTCCAACTGACTCTGGTTAGAT 1140
 QY 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGCCTCTTTAAGCTGAAACCGATGTGCACTTTCCA 1200
 DB 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGCCTCTTTAAGCTGAAACCGATGTGCACTTTCCA 1200
 QY 1201 TTTAATGATTTATGCTTGTGTTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
 DB 1201 TTTAATGATTTATGCTTGTGTTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
 QY 1261 GTGACCTGACCACTCTCAGGCACACTTTCTCAGGACCTCTCTGAGGTGTATCCAGGCCAT 1320
 DB 1261 GTGACCTGACCACTCTCAGGCACACTTTCTCAGGACCTCTCTGAGGTGTATCCAGGCCAT 1320
 QY 1321 GGTAAAGTCAATGTTGGCTCAGAACTCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTGG 1380
 DB 1321 GGTAAAGTCAATGTTGGCTCAGAACTCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTGG 1380
 QY 1381 TTACCAATGAAGTCTCCACAAATATATGTCCAAAGATCTTCAATTTCCAAAGCCTCTCACC 1440
 DB 1381 TTACCAATGAAGTCTCCACAAATATATGTCCAAAGATCTTCAATTTCCAAAGCCTCTCACC 1440
 QY 1441 AATTTCAAATGCGCAACATCTCCCACTCCCAATTTACCTTATTTCTCTTTGAGGTGTATCTA 1500
 DB 1441 AATTTCAAATGCGCAACATCTCCCACTCCCAATTTACCTTATTTCTCTTTGAGGTGTATCTA 1500
 QY 1501 CTCATTAATGCTGTGAAGCCAGTGCAGACACCTTTGTCTTAACTGACATTTTACCTCAAT 1560
 DB 1501 CTCATTAATGCTGTGAAGCCAGTGCAGACACCTTTGTCTTAACTGACATTTTACCTCAAT 1560
 QY 1561 TTTTCTTTTCTATGCTGAGTATTTTGGCATATATTTTGGCATATTTTGGCATATTTTGGCAT 1620
 DB 1561 TTTTCTTTTCTATGCTGAGTATTTTGGCATATATTTTGGCATATTTTGGCATATTTTGGCAT 1620
 QY 1621 TAATAGTTTTTGCATTTTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAA 1680
 DB 1621 TAATAGTTTTTGCATTTTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAA 1680
 QY 1681 AAGATGAATGAAGCATATATATTTTCAATTTTTTCAATTTTTTCTAGTCAACAGAGAAATC 1740
 DB 1681 AAGATGAATGAAGCATATATATTTTCAATTTTTTCAATTTTTTCTAGTCAACAGAGAAATC 1740
 QY 1741 GAAGGATTTCTGTTCAATATTTAGTAAATAATTTGAAATAAATTTGCTGCTTATATTTGTTT 1800

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2004, 12:17:23 ; Search time 1654 Seconds
(without alignments)
11939.735 Million cell updates/sec

Title: US-10-005-907-1
Perfect score: 3762
Sequence: 1 gagaaacggagctactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues 8269772

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 23Sep04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2002bs: *
8: Geneseqn2003as: *
9: Geneseqn2003bs: *
10: Geneseqn2003cs: *
11: Geneseqn2003ds: *
12: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3762	100.0	3762	6 ABN81319	Abn81319 Human mas
2	1068	28.4	1128	5 ABV20298	Abv20298 Human pro
3	1068	28.4	1128	5 ABV26130	Abv26130 Human pro
4	1068	28.4	1128	5 ABV26458	Abv26458 Human pro
5	1068	28.4	1128	5 ABV20617	Abv20617 Human pro
6	1068	28.4	1128	5 ABV20806	Abv20806 Human pro
7	1068	28.4	1128	5 ABV20439	Abv20439 Human pro
8	1068	28.4	1128	5 ABV26275	Abv26275 Human pro
9	1068	28.4	1128	5 ABV26653	Abv26653 Human pro
10	506	13.5	667	10 ADC30760	Adc30760 Human nov
11	394.2	10.5	433	5 ABV01270	Abv01270 Human pro
12	393.8	10.5	397	5 ABV10439	Abv10439 Human pro
13	392.2	10.4	421	5 ABV31276	Abv31276 Human pro
14	392.2	10.4	437	5 ABV31609	Abv31609 Human pro
15	392.2	10.4	437	5 ABV40245	Abv40245 Human pro
16	392.2	10.4	437	5 ABV40577	Abv40577 Human pro
17	391.8	10.4	404	5 ABV10104	Abv10104 Human pro
18	375	10.0	446	5 ABV00935	Abv00935 Human pro
19	342.8	9.1	359	5 ABV09593	Abv09593 Human pro
20	342.2	9.1	400	5 ABV00424	Abv00424 Human pro
21	342.2	9.1	401	5 ABV39737	Abv39737 Human pro

22	342.2	9.1	401	5	ABV30769	Abv30769 Human pro
23	342.2	9.1	401	5	ABV39958	Abv39958 Human pro
24	272.8	7.3	425	5	ABV09816	Abv09816 Human pro
25	244.2	6.5	296	5	ABV30990	Abv30990 Human pro
26	239.8	6.4	3337	4	AAK70219	Aak70219 Human imm
27	230.4	6.1	1982	5	ABV23873	Abv23873 Human pro
28	230.4	6.1	1982	5	ABV23839	Abv23839 Human pro
29	230.4	6.1	1982	5	ABV29715	Abv29715 Human pro
30	230.4	6.1	1982	5	ABV29752	Abv29752 Human pro
31	230.4	6.1	1982	5	ABV24256	Abv24256 Human pro
32	201.2	5.3	12278	5	ABA14585	Abal4585 Human ner
33	201.2	5.3	12278	5	ABA14664	Abal4664 Human ner
34	200.4	5.3	17239	4	AA336826	Aas36826 Human car
35	200.4	5.3	17239	10	ADE47520	Ade47520 Human car
36	200.4	5.3	17239	10	ABZ67711	Abz67711 Human sec
37	199.8	5.3	181257	12	ADF69677	Adf69677 Human SLC
38	199.4	5.3	2646	4	AA161012	Aa161012 Human pol
39	197.4	5.2	1743	4	AA162586	Aa162586 Human bre
40	197.4	5.2	1743	4	AA103368	Aa103368 Human rep
41	197.4	5.2	1746	4	AA162587	Aa162587 Human bre
42	197.4	5.2	1746	4	AA103369	Aa103369 Human rep
43	197.4	5.2	6685	4	AA307380	Aas07380 Human gen
44	197.4	5.2	6686	6	ABS73149	Abs73149 Human CIA
45	197	5.2	166043	12	ADL08127	Adl08127 Human gen

ALIGNMENTS

RESULT 1
ABN81319

ID ABN81319 standard; CDNA; 3762 BP.

XX AC ABN81319;

XX DT 30-AUG-2002 (first entry)

XX DE Human mast cell related gene MC1 SEQ ID NO 1.

XX KW Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;

XX KW vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma;

XX KW gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 25..432

XX FT /*tag= a

XX FT /product= "MC1"

XX PN WO200246389-A2.

XX XX 13-JUN-2002.

XX PF 07-DEC-2001; 2001WO-US046180.

XX PR 08-DEC-2000; 2000US-0251935P.

XX PR 14-MAR-2001; 2001US-0275479P.

XX PR 28-MAR-2001; 2001US-0279115P.

XX PR 02-APR-2001; 2001US-0280143P.

XX XX (UNIO) UCB SA.

XX XX Nocka K, Pirozzi G, Einstein R;

XX PI WPI; 2002-508560/54.

XX DR P-PSDB; ABB77569.

XX XX Novel isolated nucleic acids that are differentially expressed in mast

XX PT cells in patients with allergic hypersensitivity, encoding proteins

XX PT associated with mast cell regranulation and allergic hypersensitivity.

XX PS Claim 1; Page 95-97; 119pp; English.

This Page Blank (uspto)

3408 AAGTTCACAGTGTCTCTTCCCTGAGTTTACACAGTCTCTGGCAACC-AATGATCTGCT 3466
 Db |||||
 355 AAGTTCACAGTGTCTCTTCCCTGAGTTTACACAGTCTCTGGCAACC-AATGATCTGCT 414
 Qy |||||
 3467 TCGTATAATTAACAGTCTCTGAGTATTTGTAGCAATGTACCCCTTTCCATATTTATTTTG 3526
 Db |||||
 415 TCGTATAATTAACAGTCTCTGAGTATTTGTAGCAATGTACCCCTTTCCATATTTATTTTG 474
 Qy |||||
 3527 TGTGTGTAAGGCTCTCTTTAGTCAATTAATAATTTTGGAGATTCATCTATGTTTAATGTT 3586
 Db |||||
 475 GTGTGTGTAAGGCTCTCTTTAGTCAATTAATA-TGTGTGAGATTCATCTATGTTTAATG-T 532
 Qy |||||
 3587 CTATCAGTATGTTGATCTCTTACTGTCTCAGCATATCAC-CATATAGATATATATATAT 3645
 Db |||||
 533 CTATCAGTATGTTGATCTCTTACTGTCTCAGCATATCACATATAGATTTACGTATAAT 592
 Qy |||||
 3646 TTGTTAACTAATCACTCATGATGATATGTAGGATATTT--AAGTTTTTGACATATATGAATA 3703
 Db |||||
 593 GTGTGTACTCTAATCACTCATGATGATATGTAGGATATTTGACAGTGTGTTGACATATATGAATA 652
 Qy |||||
 3704 AAGTGGCTATTAATGAAAAAAGAAAAAAGAAAAAAGAAAAA 3743
 Db |||||
 653 AGTGGCTTTAATAACACAAAAAAGAAAAAAGAAAAAAGAAAAA 692

RESULT 15
 AI766257/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 456)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1086 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 437.
 Location/Qualifiers
 1..456
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2385942"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Kid11"
 /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
 plasmid DNA from the normalized library NCI CGAP Kid3 was
 prepared, and ss circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (cloneIDs 1322376-1323911, 1456007-1456775, and
 1500552-1502855). Subtraction by Bento Soares and M.

FEATURES
 source

ORIGIN

Query Match 11.9%; Score 446.6; DB 1; Length 456;
 Best Local Similarity 99.1%; Pred. No. 5.1e-53;
 Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 3283 TCAAGTGTGATTATATACAGAAAAATGTAAACACCTGTGAAGGGTAGAGTTATAAGAAATTT 3342
 Db |||||
 456 TCAAGTGTGATTATATACAGAAAAATGTAAACACCTGTGAAGGGTAGAGTTATAAGAAATTT 397
 Qy |||||
 3343 GTCAAATGATTTCACCCCATGTAGTCACTCTTATTAAGAGACAGACACGATACATCTC 3402
 Db |||||
 396 GTCAAATGATTTCACCCCATGTAGTCACTCTTATTAAGAGACAGACACGATACATCTC 337
 Qy |||||
 3403 CCAGAAAGTTCACAGAGTGTCTCTTTCCCTGAGTTTCCACAGTCTCTGGCAACCAATGATC 3462
 Db |||||
 336 CCAGAAAGTTCACAGAGTGTCTCTTTCCCTGAGTTTCCACAGTCTCTGGCAACCAATGATC 277
 Qy |||||
 3463 TGCCTTCGTATTAATAATAACTGTTCTAGATATTTGTAGCAATGTACCCCTTCCATATTTAT 3522
 Db |||||
 276 TGCCTTCGTATTAATAATAACTGTTCTAGATATTTGTAGCAATGTACCCCTTCCATATTTAT 217
 Qy |||||
 3523 TTTGTGTGTGAAGGCTCTTTTGTAGTCAATTAATAATTTTTCAGATTGATCTATGTTTAA 3582
 Db |||||
 216 TTTGTGTGTGAAGGCTCTTTTGTAGTCAATTAATAATTTTTCAGATTGATCTATGTTTAA 157
 Qy |||||
 3583 TGTCTCTATCAGTAGTTGTACATCTTACTTGTCTCAGCATATCACCATATAGATATATCTAT 3642
 Db |||||
 156 TGTCTCTATCAGTAGTTGTACATCTTACTTGTCTCAGCATATCACCATATAGATATATCTAT 97
 Qy |||||
 3643 AATTTGTTAATCTAATCACTGATGGATATGTAGGATATTTAAAGTTTTTTCACATTTATGAAT 3702
 Db |||||
 96 AATTTGTTAATCTAATCACTGATGGATATGTAGGATATTTAAAGTTTTTTCACATTTATGAAT 37
 Qy |||||
 3703 AAAGTGGCTATTAATAAGAAAAAAGAAAAAAGAAAAA 3735
 Db |||||
 36 AAAGTGGCTATTAATAAGAAAAAAGAAAAAAGAAAAA 4

Search completed: November 18, 2004, 21:26:32
 Job time : 10884 secs

120 TTAGACAACTACTGTACAAATGAAATAAATTTCTTAATCCCTGACTAACTGAATGAC 179
761 CCTCTTCTAGGCCAAGAGACCTCAGATGAACCTGAAGACTGAATTTCTGGCCATGATAG 820
180 CCTCTTCTAGGCCAAGAGACCTCAGATGAACCTGAAGACTGAATTTCTGGCCATGATAG 239
821 GAAGGGAGGTGAGACACACCTTGTATACCCCTTCCCTTTGGAGTTTATGACCAAGTGA 880
240 GAAGGGAGGTGAGACACACCTTGTATACCCCTTCCCTTTGGAGTTTATGACCAAGTGA 298
881 CCAGGATGAGTATGAAGACTGTGAATGAAATAGACTGATTTGGCAATGAAGAGTCCCAATCC 940
299 CCAGGATGAGTATGAAGACTGTGAATGAAATAGACTGATTTGGCAATGAAGAGTCCCAATCC 358
941 AACCTGACTCTGGTGTAGATCAACACTCTCTGAGGAGTTCATCTATGAGACTTTGCT 1000
359 AACCTGACTCTGGTGTAGATCAACACTCTCTGAGGAGTTCATCTATGAGACTTTGCT 418
1001 ACATAACAGAGACCTTGGTTTCCACAAACCCCTTTATTTAGCTAAAGCAATCTTTTCTAC 1060
419 ACATAACAGAGACCTTGGTTTCCACAAACCCCTTTATTTAGCTAAAGCAATCTTTTCTAC 478
1061 TGACTTCTTAAGTCTTTAGACA--AAGCTTAACTCTTTCAACCAAT--GCCAATCAGACA 1117
479 TGACTTCTTAAGTCTTTAGACAAGAGCTTTAACTCTTTCAACCAATTTGGCCAAATCAGACA 538
1118 AACTTTGAA--TCTACCTATGACC--TGTAAGCTCTCTCTGCTTCAAGATCTTGCCTC-- 1172
539 AACTTTGAA--TCTACCTATGACC--TGTAAGCTCTCTCTGCTTCAAGATCTTGCCTC 598
1173 -TTTAAGCTGAA--CCGATGTGACCTT--CCATTTAATGATTTATGCTTTGCTGTA 1226
599 TTTAAGCTGAA--CCGATGTGACCTT--CCATTTAATGATTTATGCTTTGCTGTA 658
1227 A--CTCCTGTCTCCCTAAATGATATAAAGTAAAC--GGTACCTGACCACTCAGGCACA 1283
659 AACTCTCTGCTCTGCTAAATGTTTAAAGTCAACGGGTGACCTGACCACTCAGGCACA 718
1284 CTTTCTCA--GGACCTCTGAGAGTGTATCCGAGC 1317
719 CTTTCTCAAGAACTCTCTGGAAGTGTATCCGAGC 754

RESULT 12
AW65849/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

AW65849 504 bp mRNA linear EST 06-APR-2000
hi94f08.x1 Soares NFL T.GBC S1 Homo sapiens cDNA clone
IMAGE:2979975 3', mRNA sequence.
AW65849
AW65849.1 GI:7458398
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 504)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 451.
Location/Qualifiers
1..504
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2979975"
/lab_host="DH10B"

/clone lib="Soares NFL T.GBC S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBLH19W, testis NHT, and B-cell
NCI CGAP GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
1.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 13.3%; Score 500.8; DB 2; Length 504;
Best Local Similarity 99.6%; Pred. No. 1.3e-60;
Matches 502; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

3219 CTTATCTTTTATGTCATATAAATATTTGAACATTTTACATTTGTTTATTTTAAATCAGTT 3278
Db |||||
504 CTTATCTTTTATGTCATATAAATATTTGAACATTTTACATTTGTTTATTTTAAATCAGTT 445
3279 TTACTCAAGTGTATATATACAGAAATGTAACACCTGTAAGGTAGAGTATAGAA 3338
Db |||||
444 TTACTCAAGTGTATATATACAGAAATGTAACACCTGTAAGGTAGAGTATAGAA 385
3339 TTTTGTCAATGTATTTTCAACCTGATGTCACCTCTTATGAAGAGACAGAACAGTACAT 3398
Db |||||
384 TTTTGTCAATGTATTTTCAACCTGATGTCACCTCTTATGAAGAGACAGAACAGTACAT 325
3399 CTTCCAGAAAGTTTCCACAGTGTCTCTTTTCCCTGAGTTTCCACAGTCTGCGCAACCAAT 3458
Db |||||
324 CTTCCAGAAAGTTTCCACAGTGTCTCTTTTCCCTGAGTTTCCACAGTCTGCGCAACCAAT 265
3459 GATCTCTCTGATATATATTAATCTGTTCTAGATATTTTGTAGCAATGTACCTTTCCATAT 3518
Db |||||
264 GATCTCTCTGATATATTAATCTGTTCTAGATATTTTGTAGCAATGTACCTTTCCATAT 205
3519 TTATTTTGTGTGTAAGGCTTCTTTTGTAGTATTATAAATTTTGTAGATTTTGTAGATTTATGT 3578
Db |||||
204 TTATTTTGTGTGTAAGGCTTCTTTTGTAGTATTATAAATTTTGTAGATTTTGTAGATTTATGT 145
3579 TTAATGTTCTATCAGTATTTGTATCTTTTCTGATCTTCTGATATATTTTGTAGATTTATGTAT 3638
Db |||||
144 TTAATGTTCTATCAGTATTTGTATCTTTTCTGATCTTCTGATATATTTTGTAGATTTATGTAT 85
3639 CTATAATTTGTTAACTCTAATCACTGATGATGATGATGATGATGATGATGATGATGATGAT 3698
Db |||||
84 CTATAATTTGTTAACTCTAATCACTGATGATGATGATGATGATGATGATGATGATGATGAT 25
3699 GAATAAGTGGCTATAAATGAAAA 3722
Db |||||
24 GAATAAGTGGCAATAAATGAAAA 1

RESULT 13
BZ600841
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

BZ600841 757 bp DNA linear GSS 08-JUN-2003
WHAAJ89TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
sapiens genomic clone MCF7_1-3010, genomic survey sequence.
BZ600841
BZ600841.1 GI:31509303
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 757)
Volik, S., Zhao, S., Chin, K., Brebner, J.H., Herndon, D.R., Tao, Q.,
Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P.,
Gray, J.W., and Collins, C. Sequence-based analysis of aberrant genomes
End-sequence profiling: Sequence-based analysis of aberrant genomes

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 527)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM993 row: a column: 01
High quality sequence stop: 518.
FEATURES
source
1..527
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4109064"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_55"
/notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site 1: SfII (ggcccttcggcc); Site 2: SfII
(ggcatatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CAGGCGCATTTATGGCC-3' and
3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGCGCATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

```

Query Match 13.5%; Score 506; DB 2; Length 527;
Best Local Similarity 100.0%; Pred. No. 2.3e-61;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAAACCGAGTCACTGTGAAAGATCGGAAATTTATCTCTCGCAAACTCAGTTGCTG 60
DB 21 GAGAAACCGAGTCACTGTGAAAGATCGGAAATTTATCTCTCGCAAACTCAGTTGCTG 80
QY 61 GGAGAGAATCAAAAGAGCCCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 120
DB 81 GGAGAGAATCAAAAGAGCCCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 140
QY 121 ATGACTACATTTCAAGAAAGAAAGTCAAGATCAAGATCAAGAAAGAAAGAAAGAAAGTTCATCC 180
DB 141 ATGACTACATTTCAAGAAAGAAAGTCAAGATCAAGATCAAGAAAGAAAGAAAGTTCATCC 200
QY 181 ACTTCTAATCAGGAAACGAGATGCGAGTGTCTGGAAGTGTCTACACTGTCATT 240
DB 201 ACTTCTAATCAGGAAACGAGATGCGAGTGTCTGGAAGTGTCTACACTGTCATT 260
QY 241 AATCATATCCCCATCAGAGATCCTCCCTGAGCTCCCAATGATGATGGCTATGAGACATT 300
DB 261 AATCATATCCCCATCAGAGATCCTCCCTGAGCTCCCAATGATGATGGCTATGAGACATT 320
QY 301 GACTCCCTCAGAGAAAGTGAGACAGTTTAGAGAAAGTTCAGAGACAGAAATATGCCCTT 360
DB 321 GACTCCCTCAGAGAAAGTGAGACAGTTTAGAGAAAGTTCAGAGACAGAAATATGCCCTT 380
QY 361 CTTAGGACTTCTGTTAGTAGGCTTCTCTCGACCCATGAGCATGATTAAGAGTTTG 420
DB 381 CTTAGGACTTCTGTTAGTAGGCTTCTCTCGACCCATGAGCATGATTAAGAGTTTG 440
```

```

QY 421 TTTCCACACTAAATCCTCAAGCTGCTTTATCAGCTTCCAGCAATGAAGACAATGCAGAA 480
DB 441 TTTCCACACTAAATCCTCAAGCTGCTTTATCAGCTTCCAGCAATGAAGACAATGCAGAA 500
QY 481 TAGCAGACTCTGGGGAAGTGTGTTTCCAC 506
DB 501 TAGCAGACTCTGGGGAAGTGTGTTTCCAC 526

RESULT 11
BF240893 809 bp mRNA linear EST 14-NOV-2000
60187544AF2 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4091963 5',
mRNA sequence.
BF240893
BF240893.1 GI:11154817
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 809)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM948 row: h column: 12
High quality sequence stop: 651.
FEATURES
source
1..809
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4091963"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_55"
/notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site 1: SfII (ggcccttcggcc); Site 2: SfII
(ggcatatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CAGGCGCATTTATGGCC-3' and
3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCGGCGCATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

```

Query Match 13.3%; Score 502; DB 2; Length 809;
Best Local Similarity 90.5%; Pred. No. 7.4e-61;
Matches 684; Conservative 0; Mismatches 50; Indels 22; Gaps 13;

QY 582 CTGATATGTCATCTCTGTGGCTTAGGTGAAATCATAGAAATTCACAAATGACCTAAA 641
DB 1 CTGATATGTCATCTCTGTGGCTTAGGTGAAATCATAGAAATTCACAAATGACCTAAA 60
QY 642 ATATTCTATGTTTGTGTTTAAAGTTTGAAGCATGGAGGTGATATAAAAACTTTC 701
DB 61 ATATTCTATGTTTGTGTTTAAAGTTTGAAGCATGGAGGTGATATAAAAACTTTC 119
QY 702 TTAGGACAAATATGA-BAATGAAATATAATTTCTTAATCCCTGACTGAATGATGAC 760
```



```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 728)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaabs-remail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM210 row: n column: 16
High quality sequence stop: 504.
FEATURES
source
    Location/Qualifiers
        1..728
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:30414375"
            /tissue_type="Pooled"
            /lab_host="DH10B (TI phage-resistant)"
            /clone_lib="NIH_MGC_191"
            /notes="Vector: pDNR-Lib; Site 1: Sfil (ggccattatggcc);
            Site 2: Sfil (ggccgctggcc); Library is oligo-dT primed
            and directionally cloned. PBMC - Peripheral Blood
            Mononuclear Cells. RNA was pooled from 3/6hour stimulation
            with PMA adn ionomycin. 5' and 3' adaptors were used in
            cloning as follows: 5' adaptor sequence:
            5'-CAGCGCCATTATGGCC-3' and 3' adaptor sequence:
            5'-ATTCTAGAGCCGAGCGCGGACATG-dt(30)BN-3' (where B = A,
            C, or G and N = A, C, G, or T). Average insert size 1.69
            kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
            by PCR. This library was enriched for full-length clones
            and was constructed by Clontech Laboratories (Palo Alto,
            CA). Note: this is a NIH_MGC Library."
ORIGIN
Query Match 16.3%; Score 612.2; DB 6; Length 728;
Best Local Similarity 98.4%; Pred. No. 2.8e-76;
Matches 617; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1835 ATCTTACCGAAGGTGGATGTGTAGTTCTTCCCAAGAGTTTTTAATGATATGCCAGCTTCCTAAT 1954
Db 3 ATCTTACCGAAGGTGGATGTGTAGTTCTTCCCAAGAGTTTTTAATGATATGCCAGCTTCCTAAT 1954
QY 1895 ATCCATCTAATATGCTTTCTTCCCAAGAGTTTTTAATGATATGCCAGCTTCCTAAT 1954
Db 63 ATCCATCTAATATGCTTTCTTCCCAAGAGTTTTTAATGATATGCCAGCTTCCTAAT 122
QY 1955 TTGGAGACAAAAGCCCTTAATTTGCAATGCAATTCATTATATATTTTGTATGTTACAG 2014
Db 123 TTGGAGACAAAAGCCCTTAATTTGCAATGCAATTCATTATATATTTTGTATGTTACAG 182
QY 2015 TATACAGATTGAGTATCCCTTAGATGAGATGCTTGGGACAGAGTGTGGATTTTGGATTTTCCAG 2074
Db 183 TATACAGATTGAGTATCCCTTAGATGAGATGCTTGGGACAGAGTGTGGATTTTGGATTTTCCAG 242
QY 2075 ATTTATTTTGGATTTTGGAAATATTTCCATACATATATGAGAGAGTTGGAAAATGGGAT 2134
Db 243 ATTTATTTTGGATTTTGGAAATATTTCCATACATATATGAGAGAGTTGGAAAATGGGAT 302
QY 2135 TCAAGTCTAATCATAAATTCACCTTATGTTGATATACACCTTATCTGAATAGCCTGAAG 2194
Db 303 TCAAGTCTAATCATAAATTCACCTTATGTTGATATACACCTTATCTGAATAGCCTGAAG 362

```

```

QY 2195 GTAATTTTATACAAATATTTTAAATAATTTTATGCTGTAACACAGAGTTTGGCACAATTGGA 2254
Db 363 GTAATTTTATACAAATATTTTAAATAATTTTATGCTGTAACACAGAGTTTGGCACAATTGGA 422
QY 2255 CCATCAGAAAGCAGAGTGTCACTATTTCAGTCAGTGTCTCAAAAGTTTCAGATGTATA 2314
Db 423 CCATCAGAAAGCAGAGTGTCACTATTTCAGTCAGTGTCTCAAAAGTTTCAGATGTATA 482
QY 2315 GCTGGTGATGAGTTTCATGCCAGTATCCGAGTACTTTGGGAAGCCCAAGACAGGTGGATC 2374
Db 483 GCTGGTGATGAGTTTCATGCCAGTATCCGAGTACTTTGGGAAGCCCAAGACAGGTGGATC 542
QY 2375 TCTTGAGCCAGAGTTTGGGCCAGTCTGCAACACACAGTGTGAGCTTCTTCTACAAA 2434
Db 543 TCTTGAGCCAGAGTTTGGGCCAGTCTGCAACACACAGTGTGAGCTTCTTCTACAAA 602
QY 2435 TAATTTAAAAAATTAGCCAGGTGGTG 2461
Db 603 TATTAATAAATTNAGCAGTGTGGTG 629

```

```

RESULT 8
AG182348 643 bp DNA linear GSS 09-JAN-2002
LOCUS Pan troglodytes DNA, clone: RP43-055101.TJ, genomic survey
DEFINITION
ACCESSION AG182348
VERSION AG182348.1 GI:16712028
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of Library RPCI-43
JOURNAL Unpublished
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpes@psc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library RPCI-43. This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES
source
    Location/Qualifiers
        1..643
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="RP43-055101.TJ"
            /sex="male"
            /cell_type="lymphocytes"
            /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN
Query Match 15.2%; Score 570.4; DB 9; Length 643;
Best Local Similarity 95.4%; Pred. No. 2.1e-70;
Matches 599; Conservative 0; Mismatches 26; Indels 3; Gaps 1;
QY 2751 GAATTTCTATGCCCTCTTGTGGTGTATTTTATGTGCGGAAGGGAACATAGAAATTTTG 2810

```

cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor sequence: 5'-ATTAGAGCGGCGGCACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

QY	821	GAAGGAGCTGAGACACACCTTGTATATACCCCTTCCCTTTGGAGTTTATGCACAAGTGA	880
Db	303	GAAGGAGCTGAGACACACCTTGTATATACCCCTTCCCTTTGGAGTTTATGCACAAGTGA	362
QY	881	CCAGGATGAGTCATAAGACTGATGAATAGACTGATTGTGCAATAAGAGTCCCAATTC	940
Db	363	CCAGGATGAGTCATAAGACTGATGAATAGACTGATTGTGCAATAAGAGTCCCAATTC	422
QY	941	AACCTGACTCTGGTGTAGATACACACCTGTCTGAGGGATTCATCTATGAGACTTTGCT	1000
Db	423	AACCTGACTCTGGTGTAGATACACACCTGTCTGAGGGATTCATCTATGAGACTTTGCT	482
QY	1001	ACATAACAGACCTTGTGTTCCACACCCCTTTATTTAGCTAAAGCATCTTTCTTAC	1060
Db	483	ACATAACAGACCTTGTGTTCCACACCCCTTTATTTAGCTAAAGCATCTTTCTTAC	542
QY	1061	TGACTTCTTAAGTCTTTAGACAAAGCTTAACTCTTTCAACCAATTCGCCAATCAGACAAC	1120
Db	543	TGACTTCTTAAGTCTTTAGACAAAGCTTAACTCTTTCAACCAATTCGCCAATCAGACAAC	602
QY	1121	TTTGAATCTACTATGACTGTGAAGCTCTCTCTGCTTCAAGATCTTGCCCTTTAAGCT	1180
Db	603	TTTGAATCTACTATGACTGTGAAGCTCTCTCTGCTTCAAGATCTTGCCCTTTAAGCT	662
QY	1181	GAACC--GATGTGACCTTCCATTTAATGATTAT-GTCTTTGCTTGTA--CTCCTGT	1234
Db	663	GAACCCGATGCGCCCTTTCCCATTTAATGATTATGCGCTTGTGTAACCTCTCTGSC	722
QY	1235	CTCCCTAAATGTATAAA 1253	
Db	723	TCCCTAAATGTATAAA 741	

RESULT 6
BU155429 786 bp mRNA linear EST 03-SBP-2002
LOCUS AGENCOURT_7979368 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:6108338
DEFINITION 5', mRNA sequence.
ACCESSION BU155429
VERSION BU155429.1 GI:22668961
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Plate: LLM2352 row: h column: 03
High quality sequence stop: 399.
Location/Qualifiers 1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6108338"
/tissue="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_55"
/notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccgctcgcc); Site 2: SfiI (ggcattatggcc); Double-stranded cDNA was prepared from

FEATURES	source
RESULT 7	
CD639433	
LOCUS	
DEFINITION	CD639433 728 bp mRNA linear EST 17-JUN-2003
AGENCOURT	14532881 NIH_MGC_191 Homo sapiens cDNA clone
IMAGE	30414375 5', mRNA sequence.
ACCESSION	CD639433
VERSION	CD639433.1 GI:31806240
KEYWORDS	EST.
SOURCE	Homo sapiens (human)

/lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_55"
 /note="Organ: Bone marrow; Vector: pDNR-LIB (Clontech);
 Site 1: SfiI (ggcgcctcgcc); Site 2: SfiI
 (ggccattatggcc); Double-stranded cDNA was prepared from
 cell line RNA. 5' and 3' adaptors were used in cloning as
 follows: 5' adaptor sequence: 5'-CACGGCCATATGGCC-3' and
 3' adaptor sequence:
 5'-ATTCTAGAGCGGCGGCGGCACATG-dT(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size
 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained
 inserts by PCR. This library was enriched for full-length
 clones and was constructed by Clontech Laboratories (Palo
 Alto, CA)."

ORIGIN

Query Match 18.4%; Score 692.4; DB 5; Length 785;
 Best Local Similarity 97.0%; Pred. No. 1.6e-87;
 Matches 738; Conservative 0; Mismatches 17; Indels 6; Gaps 3;

QY 2055 AGAAGTGTGTTGGATTTTCAGATTTATTTTGGATTTGGAAATTTCCATACATATATG 2114
 Db 3 AGAAGTGTGTTGGATTTTCAGATTTATTTTGGATTTGGAAATTTCCATACATATATG 62

QY 2115 AGAGAGTTGGAATGGGATTCAGTCTAATCATATAAAATTCACATTATGTTGATATAC 2174
 Db 63 AGAGAGTTGGAATGGGATTCAGTCTAATCATATAAAATTCACATTATGTTGATATAC 122

QY 2175 CTTATCTGAATAGCTGAAGTAATTTTATACAAATTTTAAATTAATTTATGCTGAAA 2234
 Db 123 CTTATCTGAATAGCTGAAGTAATTTTATACAAATTTTAAATTAATTTATGCTGAAA 182

QY 2235 CAGAGTTTGGCAGATTTGACCATTCAGAACAGAGTGTCAATTTCAAGTCAGTGCT 2294
 Db 183 CAGAGTTTGGCAGATTTGACCATTCAGAACAGAGTGTCAATTTCAAGTCAGTGCT 242

QY 2295 CAAAAGTTTCAGATGTTAAGCTGGTATGATGAGTTTCATGCCAGTATCCGAGTACTTTGG 2354
 Db 243 CAAAAGTTTCAGATGTTAAGCTGGTATGATGAGTTTCATGCCAGTATCCGAGTACTTTGG 302

QY 2355 GAAGCCAAAGCAGGTGATCTTTGAGCCAGGAGTTTGGCCAGACTGCACAAACAG 2414
 Db 303 GAAGCCAAAGCAGGTGATCTTTGAGCCAGGAGTTTGGCCAGACTGCACAAACAG 362

QY 2415 TGAGACCTGTTTCTACAAATATTAATAATTAAGCAGTGTGCTGTCGACACCTGTA 2474
 Db 363 TGAGACCTGTTTCTACAAATATTAATAATTAAGCAGTGTGCTGTCGACACCTGTA 422

QY 2475 GTCCAGGTACTCAGGAGGCTGAGGTAGTAGTGGATTTGTCAGACTGGGAGTTGAGGCTG 2534
 Db 423 GTCCAGGTACTCAGGAGGCTGAGGTAGTAGTGGATTTGTCAGACTGGGAGTTGAGGCTG 482

QY 2535 AACTGAGCCAGGATCTTGCACACATTCAGCTTGGGCAACAGAGTGAACCTGTCTC 2594
 Db 483 AACTGAGCCAGGATCTTGCACACATTCAGCTTGGGCAACAGAGTGAACCTGTCTC 542

QY 2595 AAAAAAAGTTTCAGATTTTGGAGCATTTTCAGATCTTCAGATTTGAGGATTTTCA 2654
 Db 543 AAAAAAAGTTTCAGATTTTGGAGCATTTTCAGATCTTCAGATTTGAGGATTTTCA 602

QY 2655 ACTGTACTGACCTTTTAGTCAATTCAGACATTAATCAATAGTGGACTCCAGATAACT 2714
 Db 603 ACTGTACTGACCTTTTAGTCAATTCAGACATTAATCAATAGTGGACTCCAGATAACT 662

QY 2715 CATTTGCTGTATACACA-TTTTGGCTCTCTATTCACAGATTTCTT-ATGCCCTCTTTTG 2771
 Db 663 CATTTGCTGTATACACA-TTTTGGCTCTCTATTCACAGATTTCTT-ATGCCCTCTTTTG 722

QY 2772 GTGATTTTAA---TGTGCGAAGGGAACATAGAAATTTT 2809
 Db 723 GTGATTTTAA---TGTGCGAAGGGAACATAGAAATTTT 763

RESULT 5
 BUI55309
 LOCUS
 DEFINITION
 AGENCOURT 7951175 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:6107719
 5' mRNA sequence.
 ACCESSION
 BUI55309
 VERSION
 BUI55309.1 GI:22668841
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (Bases 1 to 741)
 NIH-MGC http://mgi.nci.nih.gov/.
 AUTHORS
 National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE
 Unpublished (1999)
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: rs98@nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM2350 row: n column: 08
 High quality sequence stop: 512.

FEATURES
 source

Location/Qualifiers
 1..741
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6107719"
 /tissue_type="from acute myelogenous leukemia"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_55"
 /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
 Site 1: SfiI (ggcgcctcgcc); Site 2: SfiI
 (ggccattatggcc); Double-stranded cDNA was prepared from
 cell line RNA. 5' and 3' adaptors were used in cloning as
 follows: 5' adaptor sequence: 5'-CACGGCCATATGGCC-3' and
 3' adaptor sequence:
 5'-ATTCTAGAGCGGCGGCGGCACATG-dT(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size
 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained
 inserts by PCR. This library was enriched for full-length
 clones and was constructed by Clontech Laboratories (Palo
 Alto, CA)."

ORIGIN

Query Match 17.9%; Score 673; DB 5; Length 741;
 Best Local Similarity 97.2%; Pred. No. 8.4e-85;
 Matches 718; Conservative 0; Mismatches 15; Indels 6; Gaps 3;

QY 521 GAAACATCTCTTCTGGCTAAAGTTTAGAAAATTAATCTATTATATATCTTAGGCAAC 580
 Db 3 GAAACATCTCTTCTGGCTAAAGTTTAGAAAATTAATCTATTATATATCTTAGGCAAC 62

QY 581 TCTGATATCTGTCATCTCTGCTTAGTGAATCATAGAAATTCACAAATGACCAATGACCTAA 640
 Db 63 TCTGATATCTGTCATCTCTGCTTAGTGAATCATAGAAATTCACAAATGACCAATGACCTAA 122

QY 641 AATATCTATGTTTTTTTCTTGTAAAGTTTGGAGCATGGAGGTGATATAAAAAAATTTT 700
 Db 123 AATATCTATGTTTTTTTCTTGTAAAGTTTGGAGCATGGAGGTGATATAAAAAAATTTT 182

QY 701 CTTAGACAAATATGTAAATCAAAATAAATTTCTAATCCCTGACTAATCTGAATGGAC 760
 Db 183 CTTAGACAAATATGTAAATCAAAATAAATTTCTAATCCCTGACTAATCTGAATGGAC 242

QY 761 CCTCTTCTAGGCCAAGAGACCTCAGATGAACCTGAAGACTGAATTTGGCCATGATAG 820
 Db 243 CCTCTTCTAGGCCAAGAGACCTCAGATGAACCTGAAGACTGAATTTGGCCATGATAG 302

```

Db      543  TCATCTATGTTTAAATGTTCTATCAGTAGTGTGACATCTTACCTGTCCTAGCATATACCA 602
QY      3629  TATAGATATACATATAATTTGTTAAATCTAATCACTGATGATATAGGATATTTAAGTTT 3688
Db      603  TATAGATATACATATAATTTGTTAAATCTAATCACTGATGATATAGGATATTTAAGTTT 662
QY      3689  TTGACATTTATGATAAAGTGGCTATATAAATGAAAAAAGAAAAAAGAAAAAAGAAAAA 3748
Db      663  TTGACCTTATGAATAAAGTGGCTATATAAATGATAAAGCGACTACCAAAATTTGAAAAGACA 722
QY      3749  AAAAAAAGAAAA 3759
Db      723  AGCGGGGAAAA 733

RESULT 3
LOCUS   AQ390121
DEFINITION  704 bp DNA linear GSS 06-MAR-1999
            CITBI-E1-2548B21.TF CITBI-E1 Homo sapiens genomic clone 2548B21,
            genomic survey sequence.
ACCESSION  AQ390121
VERSION    AQ390121.1 GI:4361144
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 704)
AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
TITLE      Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1997)
COMMENT    Other GSSs: CITBI-E1-2548B21.TR
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.

FEATURES             location/Qualifiers
     source           1..704
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2548B21"
                     /sex="male"
                     /cell_type="sperm"
                     /clone_lib="CITBI-E1"
                     /note="Vector: pBelBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     Caltech Human BAC Library D"

ORIGIN
Query Match      18.5%; Score 694.2; DB 8; Length 704;
Best Local Similarity 99.6%; Pred. No. 9,1e-88;
Matches 696; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      803  GAATTCTGCCATGATAGAGAGGGAGGTGAGACACACCTTGTTATACCCCTTCCTTTTG 862
Db      6  GAAATTCGCCATGATAGAGAGGGAGGTGAGACACACCTTGTTATACCCCTTCCTTTTG 65
QY      863  GAGTTTATGCACAGTACCAGGATGAGTCTAAGACTGATGAATAGACTGATTGTGGC 922
Db      66  GAGTTTATGCACAGTACCAGGATGAGTCTAAGACTGATGAATAGACTGATTGTGGC 125
QY      923  AATAAGAGTCCCAATTCCAACCTGACTCTGGTGTAGATCACACTGCTGAGGATTC 982

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2004, 13:34:28 ; Search time 10878 Seconds
(without alignments)
12602.158 Million cell updates/sec

Title: US-10-005-907-1
Perfect score: 3762
Sequence: 1 gagaaacagctactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gssi:
9: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	696.8	18.5	749	5	BU852381
2	695.8	18.5	782	5	BQ441475
3	694.2	18.5	704	8	AQ390121
4	692.4	18.4	785	5	BU171872
5	673	17.9	741	5	BU155309
6	628.6	16.7	786	5	BU155429
7	612.2	16.3	728	6	CD639433
8	570.4	15.2	643	9	AG182348
9	533.8	14.2	549	2	BF242113
10	506	13.5	527	2	BF242113
11	502	13.3	809	2	BF240893
12	500.8	13.3	504	2	AW655849
13	499.6	13.3	757	8	BZ600841
14	487.6	13.0	848	2	BF243193
15	446.6	11.9	456	1	AI766257
16	410.8	10.9	549	4	BG461807
17	383.2	10.2	416	1	AA447527
18	328	8.7	386	8	AQ006627
19	260	6.9	920	8	AQ743145
20	232.4	6.2	439	8	B59695
21	226	6.0	1045	4	BM925124
22	221.8	5.9	437	8	AQ109243
23	198	5.3	475	8	B81885
24	196.2	5.2	611	6	CD704731

C 25	195.6	5.2	472	5	BM983330	BM983330 UI-CF-DUI
C 26	195	5.2	563	8	AQ591876	AQ591876 HS_5436_A
C 27	194.6	5.2	506	7	CR546941	CR546941 DKFZp470G
C 28	194.4	5.2	703	4	BI857056	BI857056 603383626
C 29	194.4	5.2	902	5	BU182600	BU182600 AGENCOURT
C 30	194.4	5.2	977	4	BG680170	BG680170 602626836
C 31	194.2	5.2	685	9	AG084541	AG084541 Pan trogl
C 32	193.8	5.2	500	2	BE779049	BE779049 601464810
C 33	192.4	5.1	275	1	AA618471	AA618471 mn27h08.8
C 34	192.4	5.1	522	8	AQ827545	AQ827545 HS_5304_A
C 35	192.2	5.1	1114	4	BM467056	BM467056 AGENCOURT
C 36	191.6	5.1	452	7	T74524	T74524 YC3C08.r1
C 37	191.6	5.1	592	2	BE393507	BE393507 601311094
C 38	191.6	5.1	604	8	B86762	B86762 RPC111-24P1
C 39	191.6	5.1	789	8	AQ740505	AQ740505 HS_5504_A
C 40	191.2	5.1	603	5	BU596567	BU596567 AGENCOURT
C 41	191	5.1	605	5	BM989036	BM989036 UI-H-DF0-
C 42	191	5.1	685	5	BM480744	BM480744 DKFZp686P
C 43	190.8	5.1	349	4	BG942531	BG942531 ax26603.x
C 44	190.6	5.1	563	7	CK902534	CK902534 im66910.x
C 45	190.6	5.1	576	5	BU078278	BU078278 im66910.y

ALIGNMENTS

RESULT 1
BU852381
LOCUS BU852381 749 bp mRNA linear EST 16-OCT-2002
DEFINITION AGENCOURT_10402421 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6617922
5', mRNA Sequence.
ACCESSION BU852381
VERSION BU852381.1 GI:24037344
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 749)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLCM2863 row: p column: 18
High quality sequence stop: 659.
Location/Qualifiers
1. 749
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6617922"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site: 1: sf11 (ggcgccctggcc); Site 2: sf11 (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Page Blank (uspto)

Sequence 21, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Krommal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
TITLE OF INVENTION: Sequences and Antibodies Thereto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H. CONTIG"
US-08-724-394A-21
Query Match 4.9%; Score 184; DB 2; Length 246240;
Best Local Similarity 71.7%; Pred. No. 1.6e-27;
Matches 241; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY 2294 TCAAAAAGTTTCAGATGTTAAAGCTGGTGTGATGCGAGTTCATGCCAGTGTGATCCGAGTACTTTG 2353
Db 29731 TAAAAAAGTTTCAGATGTTAAAGCTGGTGTGATGCGAGTTCATGCCAGTGTGATCCGAGTACTTTG 29790
QY 2354 GGAAGCCACAGAGGTGGATCTCTTGGAGCCAGAGTTTGAGCCAGAGTGTGAGCCACACACA 2413
Db 29791 GGAGGCTAAGCGAGGTGGATCTCTTGGAGCCAGAGTTTGAGCCAGAGTGTGAGCCACACATG 29850
QY 2414 GTGAGACCTCGTTTCTTCAAAATTAATTAAGCCAGAGTTTGAGCCAGAGTGTGAGCCACACACA 2413
Db 29851 GTGAACCCCATCTCTACTTAAATAATACAAAATTTACCCAGGATGGTGTGATGCGTCTGT 29910
QY 2474 AGTCCAGGTACTCAGAGGCTGAGGTAGTGTAGGATTTTGTAGACTGGAGGTTGAGGCT 2533
Db 29911 AATCCAGTACTCTGGAGGCTGAGTCAGGAGATCACTTGAACCCAGGAGGAGGTT 29970
QY 2534 GAATGAGCCAGATCTTGCCACACATTCAGCTTGGGCAACAGAGTGAGACCCGTGTCT 2593
Db 29971 GCAATGAGCCAGATCATGTGCTACTGCTACTCCAGCTGGGTGACAGATGAGACTCAATCT 30030

QY 2594 CAAAAAAGTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTT 2629
Db 30031 CAAAAAAGTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTTTCAGATTTT 30066
RESULT 14
US-08-724-394A-22
Sequence 22, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Krommal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
TITLE OF INVENTION: Sequences and Antibodies Thereto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H. CONTIG"
US-08-724-394A-22
Query Match 4.9%; Score 184; DB 2; Length 246240;
Best Local Similarity 71.7%; Pred. No. 1.6e-27;
Matches 241; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY 2294 TCAAAAAGTTTCAGATGTTAAAGCTGGTGTGATGCGAGTTCATGCCAGTGTGATCCGAGTACTTTG 2353
Db 29731 TAAAAAAGTTTCAGATGTTAAAGCTGGTGTGATGCGAGTTCATGCCAGTGTGATCCGAGTACTTTG 29790
QY 2354 GGAAGCCACAGAGGTGGATCTCTTGGAGCCAGAGTTTGAGCCAGAGTGTGAGCCACACACA 2413
Db 29791 GGAGGCTAAGCGAGGTGGATCTCTTGGAGCCAGAGTTTGAGCCAGAGTGTGAGCCACACATG 29850
QY 2414 GTGAGACCTCGTTTCTTCAAAATTAATTAAGCCAGAGTTTGAGCCAGAGTGTGAGCCACACACA 2413
Db 29851 GTGAACCCCATCTCTACTTAAATAATACAAAATTTACCCAGGATGGTGTGATGCGTCTGT 29910
QY 2474 AGTCCAGGTACTCAGAGGCTGAGGTAGTGTAGGATTTTGTAGACTGGAGGTTGAGGCT 2533

NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,975B
FILING DATE: 10-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20894/80
TELEPHONE: (716)263-1636
TELEFAX: (716)263-1600
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 26764 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 4q35
US-08-370-975B-1

Query Match 4.9%; Score 184; DB 1; Length 26764;
Best Local Similarity 74.7%; Pred. No. 9.2e-28;
Matches 257; Conservative 0; Mismatches 85; Indels 2; Gaps 2;
QY 2280 TTTCAGTCAGTGTCTCAAAAGTTTCAGATGTTAAG-CTGGTGATGCACTTATGCCAGT 2338
DB 16878 TGTCTGCTGCTGTGCAAAAGTTTCAGATTTTGGGCCAGTGCAGTAGTTCATGCTGT 16819
QY 2339 GATCCAGTACTTGGGAGCCAGACAGTGTGATCTTTGAGCCAGGATTTGAGGCC 2398
DB 16818 AATCCAGCACATTGGGAGCCAGGAGTGGGGAGATCGTTTGAGCTCAAGAGTTCAGAAC 16759
QY 2399 AGACTGCACACACAGTGTGAGACCTCGTTTCTCAAAATAATTAATAA-TTAGCCAGGTGT 2457
DB 16758 AGCTTAGGCAACATAGTGACAGCTCATCTCTGAAATATAAATAAATAGCCAGGTAT 16699
QY 2458 GGTGTGCACTGTGTAGTCCAGGTAATCAGGAGCTGAGGTAGTGTAGATTTGAGA 2517
DB 16698 GGTGGCGTGCACCTGTAGTCTAGTCTTGGGAGCTGAGGTGGGAGACTGCTTTGAGC 16639
QY 2518 CTGGGAGTGTAGGCTGAACTGAGCCAGGATCTTGCCACACATTCAGCTTGGGCAACA 2577
DB 16638 CCAGGAGTCAAGGCTGCAGTGTGTTTGCACACATGTAATCTCAGTCTGGGCAACA 16579
QY 2578 GAGTGAGCCCTGTCTCAAAATAAATAAAGTTTCAGATTTTG 2621
DB 16578 GAGCAAGACCCATCTCAGAAAAAATAAAGTTTCAGATTTTG 16535

RESULT 12
US-08-724-394A-20
Sequence 20, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Kronmal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.

APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
TITLE OF INVENTION: Sequences and Antibodies Thereto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..246240
OTHER INFORMATION: /note="HLA-H. CONTIG"
US-08-724-394A-20

Query Match 4.9%; Score 184; DB 2; Length 246240;
Best Local Similarity 71.7%; Pred. No. 1.6e-27;
Matches 241; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
QY 2294 TCAAAAAGTTTCAGATGTTAAGCTGGTGTAGTGCAGTTCATGCCAGTCCGAGTACTTTG 2353
DB 29731 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 29790
QY 2354 GGAAGCCAGACAGAGTGTGATCTTTGAGCCAGGAGTTTGGGCCAGACTGCAACACACA 2413
DB 29791 GGAGGTAAGCAGAGTGTGATCTTGGTGTAGGAGTTTGGAGCCAGCTGGTCAACATG 29850
QY 2414 GTGAGACCTCGTTTCTCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2473
DB 29851 GTGAAACCCCTCTCTACTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 29910
QY 2474 AGTCCAGTACTCAGGAGCTGAGTAGTAGTGTGTTGAGCTGGGAGTTGAGGCT 2533
DB 29911 AATCCAGTACTCTGGGAGGCTGAGTCAGGAGAAATCACTTGAACCCAGGAGGAGGTT 29970
QY 2534 GAAGTGCAGCAGGATCTTCCACCACTTCCAGCTTGGGCAACAGAGTGTGAGACCCCTGTCT 2593
DB 29971 GCATGAGCCAGATCATGTCTACTTCCAGCTGGGTGACAGATGAGACTCAATCT 30030
QY 2594 CAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2629
DB 30031 CAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 30066

RESULT 13
US-08-724-394A-21

Db 143 CTCAGGAGTTCAAGAAACACCTCGCCAAACATGGTGAACCCCATCTCTACTAAATAACA 202
QY 2442 AAAATATGCCAGGTGCTGCTGGTGCACACTGTAGTCCAGGTACTCAGGAGGCTGAGGTA 2501
Db 203 AAAATATGCTGGGCATGCTGGTGGGAGGCTGTAATCCCAGTTACTTTGGAGGCTAAGGTA 262
QY 2502 GTAGGATTTGTTGAGACTGGGAGGTTGAGGCTGAACTGAGCCAGGATCTTCCACCAT 2561
Db 263 GCAGAAATTTGATCTGGGAGGAGAGGTTGAGTGAGGCAAGATTGCCCTCTGCAC 322
QY 2562 TCCAGCTGGGCAACAGAGTGAGACCCCTGTCTCAAAAAAAGTTTCAGATTT 2619
Db 323 TCCAGCTGGGTGACAGAGTCAGACTCTGTCTCAAAAAAAGTTTCAGATTT 380

RESULT 9

US-09-754-250-3/c
; Sequence 3, Application US/09754250
; Patent No. 6376225
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO01063
; CURRENT APPLICATION NUMBER: US/09/754,250
; CURRENT FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(111282)
; OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

Query Match 4.9%; Score 184.2; DB 3; Length 111282;
Best Local Similarity 76.8%; Pred. No. 1.2e-27;
Matches 225; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 2316 CTGCTGATGAGTTTCATGCCAGTGCATCCGAGTACTTTGGGAAGCCAGACAGGTGGATCT 2375
Db 14773 CGGCTGAGTGGCTCATGCTGTAATCCAGTACTTTGGGAGGCCAAGGTGGGTGAACA 14714
QY 2376 CTTGAGCCCGAGGTTTGGGCCAGACTGCACACACAGTGAGACCTCTGTTCTACAAAT 2435
Db 14713 TTTGAGGTCAGGAGTTCAAGACACAGCCCTGGCCAACTGGTGAAACCCCATCTCTACTAA 14654
QY 2436 AATTAAAAAATTAGCCAGGTGGTGGTGCACACCTGTAGTCCAGGTACTCAGGAGGCT 2495
Db 14653 AATACAAAAATTATGCAAGGTGGTGGTGGTGCATGCTGTAGTCCAGGTACTCAGGAGGCT 14594
QY 2496 GAGGTAGTAGGATTTGAGACTGGGAGGTTGAGGCTGAACTGAGCCAGGATCTTGCCA 2555
Db 14593 GAGGAGGAGAAATAGCTTGAGCTGGATCGGAGGTTGGGTGAGCCGAGATCAGCCA 14534
QY 2556 CCACATTCAGCTTGGGCAACAGAGTGAGACCCCTGTCTCAAAAAAAGTTTCAGATTT 2608
Db 14533 CTCGACTCCAGTCTGGGTGAGAGTGAGACCCCTGTCTCAAAACTAACTAA 14481

RESULT 10

US-08-370-975B-6/c
; Sequence 6, Application US/08370975B
; Patent No. 5622851
; GENERAL INFORMATION:
; APPLICANT: Maley, Frank
; APPLICANT: Maley, Gladys F.
; APPLICANT: Weiner, Karen X.B.
; TITLE OF INVENTION: Human Deoxycytidylate Deaminase Gene

; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,975B
; FILING DATE: 10-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 20894/80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716)263-1636
; TELEFAX: (716)263-1600
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20303 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 4q35
US-08-370-975B-6

Query Match 4.9%; Score 184; DB 1; Length 20303;
Best Local Similarity 74.7%; Pred. No. 8.6e-28;
Matches 257; Conservative 0; Mismatches 85; Indels 2; Gaps 2;

QY 2280 TTTCAAGTCAGTCTCAAAAGTTTCAGATGTTAAG-CTGGTATGCAGTTTCATGCCAGT 2338
Db 14915 TGTATGCTGCTGTGCANAAGTTTCAGATTTGGCCAGGTGCAGTAGTTTCATGCTGT 14856
QY 2339 GATCCAGTACTTTGGGAAGCCAGACAGTGGATCTCTTGAGCCAGGAGTTTGAGGCC 2398
Db 14855 AATCCAGACATTTGGGAGGCCAAGTGGGGAGATCGTTTGAGCTCAAGAGTTCAAGAAC 14796
QY 2399 AGACTCCACACACAGTGAGACCTCGTTTCTACAAATAATTAATAA- TTAGCCAGGTGT 2457
Db 14795 AGCTTAGGCAACATAGTGACAGCTCATCTCTACTGAAAAATAAAAAAAGTTAGCCAGGTAT 14736
QY 2458 GGTGGTGACACCTGTAGTCCCAGGTACTCAGGAGGCTGAGGTAGGATGTTTGGAGA 2517
Db 14735 GGTGGGTGCACCTGTAGTCTTCTAGTACTTGGAGGCTCAGGTGGAGGACTGCTTGAGC 14676
QY 2518 CTGGGAGGTTGAGGCTGAACCTGAGCCAGGATCTTGCCACCATTCACAGCTTGGGCAACA 2577
Db 14675 CCAGGAGGTCAGAGGCTGAGTGCAGTGTGTTTGCCACCAGTGTACTCCAGTCTGGGCAACA 14616
QY 2578 GAGTGAGACCTGTCTCAAAAAAAGTTTCAGATTTTG 2621
Db 14615 GAGCAAGACCCCATCTCAGAAAAAAGTTTCAGATTTTG 14572

RESULT 11

US-08-370-975B-1/c
; Sequence 1, Application US/08370975B
; Patent No. 5622851
; GENERAL INFORMATION:
; APPLICANT: Maley, Frank
; APPLICANT: Maley, Gladys F.
; APPLICANT: Weiner, Karen X.B.
; TITLE OF INVENTION: Human Deoxycytidylate Deaminase Gene

LOCATION: 15242..15260
OTHER INFORMATION: 17-41-250.mis complement
NAME/KEY: primer bind
LOCATION: 42199..42217
OTHER INFORMATION: 20-841-149.mis
NAME/KEY: primer bind
LOCATION: 42219..42237
OTHER INFORMATION: 20-841-149.mis complement
NAME/KEY: primer bind
LOCATION: 45423..45441
OTHER INFORMATION: 20-842-115.mis
NAME/KEY: primer bind
LOCATION: 45443..45461
OTHER INFORMATION: 20-842-115.mis complement
NAME/KEY: primer bind
LOCATION: 77039..77057
OTHER INFORMATION: 20-853-415.mis
NAME/KEY: primer bind
LOCATION: 77059..77077
OTHER INFORMATION: 20-853-415.mis complement
NAME/KEY: misc binding
LOCATION: 1227..1251
OTHER INFORMATION: 20-828-311.probe
NAME/KEY: misc binding
LOCATION: 12335..12359
OTHER INFORMATION: 17-42-319.probe
NAME/KEY: misc binding
LOCATION: 15229..15253
OTHER INFORMATION: 17-41-250.probe
NAME/KEY: misc binding
LOCATION: 42205..42230
OTHER INFORMATION: 20-841-149.probe
NAME/KEY: misc binding
LOCATION: 45430..45454
OTHER INFORMATION: 20-842-115.probe
NAME/KEY: misc binding
LOCATION: 77045..77070
OTHER INFORMATION: 20-853-415.probe
US-09-750-580-1

Query Match 4.9%; Score 185; DB 4; Length 81001;
Best Local Similarity 77.5%; Pred. No. 7.7e-28;
Matches 224; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 2318 GGTGATGCGTTCATGCCAGTGTCCGAGTACTTTGGAGGCGCAACAGAGGTGATCTCT 2377
DB 69837 GGGCCAGTGGCTCATGCTATGATCTCTAGTCTTTGGAGGCTGAGGTGGGTAGTGTCT 69896
QY 2378 TGAGCCAGGAGTTTGAGGCCAGCTGCAACACAGTGTGAGCTCGTTTCTACAATAA 2437
DB 69897 TGAGCCAGGAGTTTGAGGCCAGCTGCAACACAGTGTGAGCTCGTTTCTACAATAA 69956
QY 2438 TTAATAAATAGCCAGGTGTGGTGTGACACACCTGTAGTCCAGGTACTCAGGAGGTGA 2497
DB 69957 TTAATAAATAGCTGGCATGTGGCCACCTGTAGTCCAGGTACTCAGGAGGTGA 70016
QY 2498 GGTAGTAGATTTTGTAGCTGGAGGTGAGCTGAATGAGCCAGGTCTTGCACC 2557
DB 70017 GGTGAAGATTCATCTGAGTCTGGAGGTGAGCTGAGCTGAGCTGTGATGAGCCACT 70076
QY 2558 ACATTCAGCTTGGGCAACAGTGTGAGCCCTGTCTCAAAAAAATAA 2606
DB 70077 GCATCCAGCGGTGACAGAGTAACTGTCTCAAAAAAATAA 70125

RESULT 7
US-09-497-855A-40
; Sequence 40, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UWO1523

CURRENT APPLICATION NUMBER: US/09/497,855A
CURRENT FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/120,592
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: 60/118,760
PRIOR FILING DATE: 1999-02-05
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.0
SEQ ID NO 40
LENGTH: 161652
TYPE: DNA
ORGANISM: Homo sapiens;
US-09-497-855A-40
Query Match 4.9%; Score 184.8; DB 4; Length 161652;
Best Local Similarity 77.1%; Pred. No. 1e-27;
Matches 225; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 2327 GTTCATGCCAGTGTCCGAGTACTTTGGAGGCCAAGACAGGTGATCTCTTGGAGGCCAG 2386
DB 154102 GCTCAGCCTATAATCTCAGCACTTTGGAGGCCAGGCGGTGGATCACTTGGAGGTGAG 154161
QY 2387 GAGTTTGGAGCCAGACTGCAACACAGTGTGAGTCTCTCAAAATAATTAATAAT 2446
DB 154162 GAAATTTGAGCCAGTCTGGCCAATACGGTGAACCTGTCTTCTAAATAATACAAAT 154221
QY 2447 TAGCCAGGTGTGGTGTGACACACCTGTAGTCCAGGTACTCAGGAGGCTGAGGTAGTAGG 2506
DB 154222 TACCAGGTGTGGTGTGACACCTGTAGTCTTAAAGTACTCAGGAGGCTGAGGCGAGAG 154281
QY 2507 ATTGTTGAGACTGGAGGTGTGAGGTGAACTGAGCCAGGATCTTCCACCAATTCAGATT 2566
DB 154282 ATTGCTTGAACCTGGGAGCAGAGGTTCAGTGTGAGCCAGGATCGCACCTGCACTCCAG 154341
QY 2567 CTGGGCAACAGAGTGTGAGCCCTGTCTCAAAAAAATAATTAATAAT 2618
DB 154342 CCTGAGTGCAAGCAAGACTCTGTCTCAAAAAAATAATTAATAAT 154393

RESULT 8
US-09-621-976-13959
; Sequence 13959, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S. J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 13959
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-13959

Query Match 4.9%; Score 184.4; DB 4; Length 399;
Best Local Similarity 69.3%; Pred. No. 2.6e-28;
Matches 248; Conservative 2; Mismatches 108; Indels 0; Gaps 0;
QY 2262 AAGCAGAGTGTCACTATTTCAGTCAGTGTCTCAAAAGTTCAGATGTTAAGTCGTGTG 2321
DB 23 AAGGAAAAAAGAAATTTTGTGTTAGTCACTCAATAAATGTTCTGTTGGCCGCG 82
QY 2322 ATGAGTTTCATGCCAGTGTCCGAGTACTTTGGAGGCCAAGACAGGTGATCTCTTGAAG 2381
DB 83 TGGTGGCTCATGCTGTAACTCTAGCACTTTGGAGGCTGAGGAGGTGATCATCTTTGAG 142
QY 2382 CCAGGAGTTTGGAGGCCAGACTGCAACACAGTGTGAGCCCTCGTTTCTACAAATAATA 2441


```

; OTHER INFORMATION: exon 8
; FEATURE:
; NAME/KEY: exon
; LOCATION: 12254..12340
; OTHER INFORMATION: exon 9
; FEATURE:
; NAME/KEY: exon
; LOCATION: 12854..13023
; OTHER INFORMATION: exon 10
; FEATURE:
; NAME/KEY: exon
; LOCATION: 13308..13429
; OTHER INFORMATION: exon 11
; FEATURE:
; NAME/KEY: exon
; LOCATION: 16567..16667
; OTHER INFORMATION: exon 12
; FEATURE:
; NAME/KEY: exon
; LOCATION: 16775..16945
; OTHER INFORMATION: exon 13
; FEATURE:
; NAME/KEY: exon
; LOCATION: 17063..17554
; OTHER INFORMATION: exon 14
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 17555..20674
; OTHER INFORMATION: 3'regulatory region
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1128
; OTHER INFORMATION: 10-508-191 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1182
; OTHER INFORMATION: 10-508-245 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1559
; OTHER INFORMATION: 10-509-284 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1570
; OTHER INFORMATION: 10-509-295 : deletion of C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1827
; OTHER INFORMATION: 10-510-173 : variable motif ATTTA or TTTTTT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2048
; OTHER INFORMATION: 10-511-62 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2323
; OTHER INFORMATION: 10-511-337 : insertion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2341
; OTHER INFORMATION: 10-512-36 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2623
; OTHER INFORMATION: 10-512-318 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2832
; OTHER INFORMATION: 10-513-250 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2844
; OTHER INFORMATION: 10-513-262 : polymorphic base C or T

```

```

; FEATURE:
; NAME/KEY: allele
; LOCATION: 2934
; OTHER INFORMATION: 10-513-352 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2947
; OTHER INFORMATION: 10-513-365 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 3802
; OTHER INFORMATION: 12-206-81 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4062
; OTHER INFORMATION: 10-343-231 : deletion of C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4088
; OTHER INFORMATION: 12-206-366 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4109
; OTHER INFORMATION: 10-343-278 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4170
; OTHER INFORMATION: 10-343-339 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 5903
; OTHER INFORMATION: 10-346-23 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6019
; OTHER INFORMATION: 10-346-141 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6141
; OTHER INFORMATION: 10-346-263 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6183
; OTHER INFORMATION: 10-346-305 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6338
; OTHER INFORMATION: 10-347-74 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6375
; OTHER INFORMATION: 10-347-111 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6429
; OTHER INFORMATION: 10-347-165 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6467
; OTHER INFORMATION: 10-347-203 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 6484
; OTHER INFORMATION: 10-347-220 : polymorphic base A or G

```

```

Query Match      4.9%; Score 185.8; DB 4; Length 20674;
Best local Similarity 73.2%; Pred. No. 3.7e-28;
Matches 238; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy      2297 AAAAGTTTCAGATGTTAAAGCTGGTGATGCAGTTCATGCCAGTGTATCCGAGTACTTTGGGA 2356
Db      534 AAAATCTGAAGAGCTGGTGGCATGTGGCTCACGCTGTATCCGACACTTTGGGA 593

```


OTHER INFORMATION: 10-347-165 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6467
OTHER INFORMATION: 10-347-203 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6484
OTHER INFORMATION: 10-347-220 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6534
OTHER INFORMATION: 10-347-271 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 6611
OTHER INFORMATION: 10-347-348 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 7668
OTHER INFORMATION: 10-348-391 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8608
OTHER INFORMATION: 10-349-47 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 8658
OTHER INFORMATION: 10-349-97 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8703
OTHER INFORMATION: 10-349-142 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 8777
OTHER INFORMATION: 10-349-216 : deletion of CTG
NAME/KEY: allele
LOCATION: 8785
OTHER INFORMATION: 10-349-224 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 8926
OTHER INFORMATION: 10-349-368 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12171
OTHER INFORMATION: 10-350-72 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12429
OTHER INFORMATION: 10-350-332 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13341
OTHER INFORMATION: 10-507-170 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 13492
OTHER INFORMATION: 10-507-321 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 13524
OTHER INFORMATION: 10-507-353 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13535

Query Match 4.9%; Score 185.8; DB 4; Length 20674;
Best Local Similarity 73.2%; Pred. No. 3.7e-28;
Matches 238; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
QY 2297 AAAAGTTTCAGATGTTAAGCTGTGTGTCAGTTCATGCCAGTCCAGTATCCGAGTACTTTGGGA 2356
Db 534 AAAAATCTGAAAGCTGGCTGGCGATGGTGGCTACGCCCTGTAAATCCAGCACTTTGGGA 593
QY 2357 AGCCAGACAGGTGGATCTTTGAGCCCGAGGAGTTTGAGCCAGACTGACACACACAGTG 2416
Db 594 AGCCAGAGTGGTGGATCATCTGAGGTCAGAGTTTAAGACCCAGCTGGCCACATGGCA 653
QY 2417 AGACCTCGTTTCTACAAATTAATTAATAATAGCCAGGTGTGTGGTGACACCTGTAGT 2476
Db 654 AAACCTCGTCTCTACTAAAAAATACAAAACTAGCCAGGTTGGTGGTGACACCTGTAA 713
QY 2477 CCCAGGTACTCAGAGGCTGAGGTAGTAGATTGTTTGACCTGGAGGTTGAGGCTGAA 2536
Db 714 CCAGCTACTCAGAGAGCTGAGCAGAGAAATCGCTTGAACCTGGGAGGTGCA 773
QY 2537 CTGAGCCAGGATCTTGCCACCAATTCAGCTTTGGGCAACAGAGTGAGACCCCTGTCTCAA 2596

Db 774 GTGAGCCGAGATCAGCCACTGCACTCTAGCTGGCAACAGAGCAAGACTCTATCTCAA 833
QY 2597 AAAAAAAGTTTCAGATTTTG 2621
Db 834 AAAAAAAGAGAGTTTCTCTG 858
RESULT 5
US-10-170-097-651
; Sequence 651, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilva
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GEN-T114XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 651
; LENGTH: 20674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1123..3123
; OTHER INFORMATION: 5'regulatory region
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3124..3297
; OTHER INFORMATION: exon 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3871..4072
; OTHER INFORMATION: exon 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5552..5633
; OTHER INFORMATION: exon 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5758..5880
; OTHER INFORMATION: exon 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5996..6099
; OTHER INFORMATION: exon 5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6349..6509
; OTHER INFORMATION: exon 6
; FEATURE:
; NAME/KEY: exon
; LOCATION: 7379..7522
; OTHER INFORMATION: exon 7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 8645..8854

APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
FILE REFERENCE: GENST.051Cp1
CURRENT APPLICATION NUMBER: US 09/641,638
CURRENT FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 651
LENGTH: 20674
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1123..3123
OTHER INFORMATION: 5'regulatory region
NAME/KEY: exon
LOCATION: 3124..3297
OTHER INFORMATION: exon 1
NAME/KEY: exon
LOCATION: 3871..4072
OTHER INFORMATION: exon 2
NAME/KEY: exon
LOCATION: 5552..5633
OTHER INFORMATION: exon 3
NAME/KEY: exon
LOCATION: 5758..5880
OTHER INFORMATION: exon 4
NAME/KEY: exon
LOCATION: 5996..6099
OTHER INFORMATION: exon 5
NAME/KEY: exon
LOCATION: 6349..6509
OTHER INFORMATION: exon 6
NAME/KEY: exon
LOCATION: 7379..7522
OTHER INFORMATION: exon 7
NAME/KEY: exon
LOCATION: 8645..8854
OTHER INFORMATION: exon 8
NAME/KEY: exon
LOCATION: 12254..12340
OTHER INFORMATION: exon 9
NAME/KEY: exon
LOCATION: 12854..13023
OTHER INFORMATION: exon 10
NAME/KEY: exon
LOCATION: 13308..13429
OTHER INFORMATION: exon 11
NAME/KEY: exon
LOCATION: 16567..16667
OTHER INFORMATION: exon 12
NAME/KEY: exon
LOCATION: 16775..16945
OTHER INFORMATION: exon 13
NAME/KEY: exon
LOCATION: 17063..17554
OTHER INFORMATION: exon 14
NAME/KEY: misc feature
LOCATION: 17555..20674
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 1128
OTHER INFORMATION: 10-508-191 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1182
OTHER INFORMATION: 10-508-245 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1559
OTHER INFORMATION: 10-509-284 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1570
OTHER INFORMATION: 10-509-295 : deletion of C
NAME/KEY: allele
LOCATION: 1827
OTHER INFORMATION: 10-510-173 : variable motif ATTGA or TTTTT
NAME/KEY: allele
LOCATION: 2048
OTHER INFORMATION: 10-511-62 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2323
OTHER INFORMATION: 10-511-337 : insertion of T
NAME/KEY: allele
LOCATION: 2341
OTHER INFORMATION: 10-512-36 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 2623
OTHER INFORMATION: 10-512-318 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2832
OTHER INFORMATION: 10-513-250 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2844
OTHER INFORMATION: 10-513-262 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2934
OTHER INFORMATION: 10-513-352 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2947
OTHER INFORMATION: 10-513-365 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 3802
OTHER INFORMATION: 12-206-81 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 4062
OTHER INFORMATION: 10-343-231 : deletion of C
NAME/KEY: allele
LOCATION: 4088
OTHER INFORMATION: 12-206-366 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4109
OTHER INFORMATION: 10-343-278 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4170
OTHER INFORMATION: 10-343-339 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 5903
OTHER INFORMATION: 10-346-23 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6019
OTHER INFORMATION: 10-346-141 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6141
OTHER INFORMATION: 10-346-263 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6183
OTHER INFORMATION: 10-346-305 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6338
OTHER INFORMATION: 10-347-74 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6375
OTHER INFORMATION: 10-347-111 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6429

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2004, 13:35:33 ; Search time 288 Seconds

(without alignments)
9284.680 Million cell updates/sec

Title: US-10-005-907-1

Perfect score: 3762
Sequence: 1 gagaaacagctactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/6C_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	187.8	5.0	1332	4	US-09-584-568C-1
2	186.6	5.0	2336	4	US-10-140-002-383
3	186.2	4.9	1426	4	US-09-620-312D-967
4	185.8	4.9	20674	4	US-09-641-638-651
5	185.8	4.9	20674	4	US-10-170-097-651
6	185.8	4.9	81001	4	US-09-750-580-1
7	184.8	4.9	161652	4	US-09-497-855A-40
8	184.4	4.9	399	4	US-09-621-976-13959
9	184.2	4.9	111282	3	US-09-754-250-3
10	184	4.9	20303	1	US-08-370-975B-6
11	184	4.9	26764	1	US-08-370-975B-1
12	184	4.9	246240	2	US-08-724-394A-20
13	184	4.9	246240	2	US-08-724-394A-21
14	184	4.9	246240	2	US-08-724-394A-22
15	183.8	4.9	2477	1	US-08-429-742-1
16	183.4	4.9	2932	4	US-09-016-434-1419
17	183.4	4.9	2932	4	US-09-054-272-5
18	183	4.9	6669	3	US-09-212-971-5
19	183	4.9	6669	3	US-08-800-929A-5
20	183	4.9	6669	3	US-09-617-053A-5
21	183	4.9	6669	4	US-09-672-717-230
22	182.8	4.9	1430	4	US-09-489-847-39
23	182.6	4.9	48763	4	US-09-916-204-3
24	182.6	4.9	48763	4	US-10-282-048-3
25	182.6	4.9	174493	4	US-09-804-471A-3
26	182.6	4.9	174493	4	US-10-238-709-3
27	182.2	4.8	1855	4	US-09-023-655-1380

c	28	182	4.8	786431	4	US-09-751-389-3	Sequence 3, Appli
	29	181.8	4.8	4768	4	US-09-526-193A-16	Sequence 16, Appli
	30	181.4	4.8	107820	4	US-09-792-616-1	Sequence 1, Appli
	31	180.6	4.8	1504	4	US-10-067-422-8	Sequence 8, Appli
	32	180.6	4.8	1519	4	US-09-716-129-45	Sequence 45, Appli
	33	180.6	4.8	44848	4	US-09-435-739-42	Sequence 42, Appli
	34	180.6	4.8	44848	4	US-09-988-113-42	Sequence 42, Appli
	35	180.6	4.8	66933	4	US-09-544-398B-11	Sequence 11, Appli
	36	180.6	4.8	66933	4	US-09-543-771-11	Sequence 11, Appli
	37	180.6	4.8	72049	4	US-09-544-398B-9	Sequence 9, Appli
	38	180.6	4.8	72049	4	US-09-543-771-9	Sequence 9, Appli
	39	180.4	4.8	72604	3	US-09-268-992-7	Sequence 7, Appli
	40	180.4	4.8	72604	3	US-09-657-474-7	Sequence 7, Appli
	41	180.4	4.8	99500	3	US-09-798-096-10	Sequence 10, Appli
	42	180	4.8	529	4	US-09-621-976-2252	Sequence 2252, Ap
	43	180	4.8	66804	4	US-09-740-041-3	Sequence 3, Appli
	44	179.6	4.8	368	4	US-09-621-976-10146	Sequence 10146, A
	45	179.6	4.8	2240	4	US-09-814-351-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-584-568C-1
; Sequence 1, Application US/09584568C
; Patent No. 6500657
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria, Alexandra et al.
; TITLE OF INVENTION: 33167, A NOVEL HUMAN HYDROLASE AND USES THEREFOR
; FILE REFERENCE: MNI-140
; CURRENT APPLICATION NUMBER: US/09/584,568C
; CURRENT FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 60/193,954
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(991)
US-09-584-568C-1

Query Match	5.0%	Score 187.8;	DB 4;	Length 1332;
Best Local Similarity	76.2%	Pred. No. 7.2e-29;		
Matches 231;	Conservative	0;	Mismatches 72;	Indels 0;
Gaps	0;			
Qy	2306	AGATGTTAAGCTGTCGTCATGCCAGTTCATGCCAGTGATCCGAGTACTTTGGAGCCAAAGAC	2365	
Db	1007	AGAAGATGCCGGCGCTGGTGGCTCATGCTGTAATTCAGCACTTTGGAGGCTAAGGC	1066	
Qy	2366	AGGTGGATCTCTTTGAGCCAGGAGTTTGAGGCCAGACTGCACAACACAGTGAGACCTCGT	2425	
Db	1067	GGGAGGATGATTTGAGCCAGGAGTTGGAGACCAGCTGCCCAACATGTTGAACCTGT	1126	
Qy	2426	TTCTACAATAATTAATAAATTTAGCCAGGTGGTGGTGACACCTGTAGTCCAGGTAC	2485	
Db	1127	CTCTACTAAATAATACAAAAATTAGCCCTGGCGTGGTGGTGACACCTGTATATCCAGGTAC	1186	
Qy	2486	TCAGGAGGCTGAGTGTAGTATTGTTGAGACTGGGAGGTTGAGGCTGAACCTGAGCCAG	2545	
Db	1187	TCGGAGGCTGAGCAGGAGAACTACTTGAACCTGGAGCAGAGGTTGCAATGAGCCGA	1246	
Qy	2546	GATCTTGGCCACCACTTCAGCTTGGGCAACAGAGTGAGACCTGTCTCAAAAAA	2605	
Db	1247	GATCACCACCTACTACCTCCAGCCTAGGCAACAGAGCAAGACTCTGTCTCAAAAAA	1306	
Qy	2606	AAA 2608		
Db	1307	AAA 1309		

This Page Blank (uspto)

us-10-005-907-2.rge

Fri Nov 19 05:43:45 2004

Search completed: November 19, 2004, 02:51:07
Job time : 3587 secs

```
Db 404 GGGAACTCTGCTGAAGACTATAGATGTTCCCTGAAAGCTGAGAGCCCGAGAG 463
Qy 102 PheArgGluArgSerGluThrGluTyrAlaLeuArgThrSerValSerArgProCys 121
Db 464 TCCTTGGGAGAACTGAGACTGAGTATTCACCTTCTACATATGCTTCTACAGACCCAGG 523
Qy 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 524 CATGCCCGATCCCGAAGAGATGAATGAACCTTCTCATGCCTCAC 568

RESULT 14
LOCUS CQ733169 2505 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 19103 from Patent WO02068579.
ACCESSION CQ733169
VERSION CQ733169.1 GI:42315682
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Hillman,J.L., Jones,A.L., Yu,J.Y., Wright,R.J., Gietzen,D., Liu,T.,
Yap,P., Dahl,C.R., Momiya,M.G., Bradley,D., Rohatgi,S.,
Harris,B., Roseberry,Ann,M., Gerstin,B.H., Peralta,C.H.,
David,M.H., Panzer,Scott,R., Flores,V., Daffo,A., Marwaha,R.,
Chen,A.J., Chang,S.C., Au,A.P. and Inman,R.R.
TITLE Secretory molecules
JOURNAL Patent: WO 02068579-A 46 14-MAR-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
Location/Qualifiers
1..2505
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.35e-07 Length: 2505
Score: 166.50 Matches: 43
Percent Similarity: 58.87% Conservative: 30
Best Local Similarity: 34.68% Mismatches: 46
Query Match: 23.29% Indels: 5
DB: 6 Gaps: 4

US-10-005-907-2 (1-135) x CQ733169 (1-2505)
Qy 14 GluAsnGlnLysLysProLysLysGlyAsnProAspGluGluArgLysGlnGluMet 33
Db 2068 GAACATGGGAAGAGGCAATTCAG-----AAGAAAGAAATGTTAAGAGGAAAAATA 2121
Qy 34 Thr-Thr-PheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerSerThr 53
Db 2122 CTCAATTTTGAAGAAGG---CAAGATTCCCAAAACGAAATGAAGAATGTCATCTACT 2178
Qy 54 SerAsnGlnGluAsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleAsn 73
Db 2179 CCATCCAGACAATGTGACACAGACCTACTCAGAGAGCTGTGCTATACCTCATCAAT 2238
Qy 74 His---IleProHisGlnArgSerSerLeuSerSerAsnAspGlyTyrGluAnile 92
Db 2239 CATCGGGTTCCTGTACAAGGCCATCAGGGAACCTGCTGCTGAAGAGTACTATGAGATGTT 2298
Qy 93 AspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeu 112
Db 2299 CCCTGCAAGCTGAGAGCCAGAGAGTCTTGGGAGGAACTGAGACTGAGTATTCACCT 2358
Qy 113 LeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAspTyrGluVal 131
Db 2359 CTACATATGCTTCTACAGACCCAGGCGATGCCCGATCCCGAAGATGAATGAACCTT 2418
Qy 132 ValPheProHis 135
Db 2419 CTCATGCCTCAC 2430
```

```
RESULT 15
LOCUS AX740457 1643 bp DNA linear PAT 10-MAY-2003
DEFINITION Sequence 46 from Patent WO02020756.
ACCESSION AX740457
VERSION AX740457.1 GI:30523621
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Stuart,J., Lincoln,S.E., Altus,C.M., Dufour,G., Chalup,M.,
Hillman,J.L., Jones,A.L., Yu,J.Y., Wright,R.J., Gietzen,D., Liu,T.,
Yap,P., Dahl,C.R., Momiya,M.G., Bradley,D., Rohatgi,S.,
Harris,B., Roseberry,Ann,M., Gerstin,B.H., Peralta,C.H.,
David,M.H., Panzer,Scott,R., Flores,V., Daffo,A., Marwaha,R.,
Chen,A.J., Chang,S.C., Au,A.P. and Inman,R.R.
TITLE Secretory molecules
JOURNAL Patent: WO 02020756-A 46 14-MAR-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
Location/Qualifiers
1..1643
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: LG:018739.2:2000SEP08"

ORIGIN
Alignment Scores:
Pred. No.: 5.22e-07 Length: 1643
Score: 159.00 Matches: 47
Percent Similarity: 50.64% Conservative: 32
Best Local Similarity: 30.13% Mismatches: 55
Query Match: 22.24% Indels: 22
DB: 6 Gaps: 6

US-10-005-907-2 (1-135) x AX740457 (1-1643)
Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db 1096 ATGGGAAATTCCTCTGCTGAGAGAAAACAGCGCGCAGCAACACTCAAGAGATGCCTGG 1155
Qy 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db 1156 AATGTGAGATGCAAAAGCCCAACAGAGAACATCCAGATGCTGGGATCACCATATCGCT 1215
Qy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 1216 GAAGGCTGTTTCTGCTTCCATGGAAAAAATACTCATTTTGAAGAAGG---CAAGAT 1272
Qy 43 GlnAspLysLysSerGlnGluValSerSerThrSer---AsnGlnGluAsnGluAsnGly 61
Db 1273 TCCAAAACGAAATGAAGAATGCTATCTATCCCATCCATCCAGCAGACAATGTTGACCAG 1332
Qy 62 SerGlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 80
Db 1333 ACCTACTCAGAGAGCTGTGCTATACCCCTCATCATTCGGGTTCCTGTACAAGGCA 1392
Qy 81 SerLeuSerSerAsnAspAspGlyTyrGluAnileAspSerLeuThrArgLysValArg 100
Db 1393 TCAGGGAACCTGCTGAAGAGTACTATGAGAATGTTCCCTGCAAGACTCAGAGACCCAGA 1452
Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuArgThrSerValSerArgPro 120
Db 1453 GAGTCCCTGGGAGAACTGAGACTGAGTATTCACCTTCTANATATGCTTCTACAGACCC 1512
Qy 121 CysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 1513 AGGCATGCCCGATCCCGAAGAGATGAATGAACCTTTCATGCCTCAC 1560
```

Medicine, Stanford University, 269 Campus Drive, CCSR building,
Room 1100, Palo Alto, CA 94305-5151, USA
Location/Qualifiers
1..1654
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
173..709
/note="germinal center-associated lymphoma; contains
modified ITAM motif, similar to mouse M17"
/codon_start=1
/product="HCAL"
/protein_id="RAO21701.1"
/db_xref="GI:27335683"
/translation="MGNSLLRNRQNTQEMPNVNRMSQKPTSRCDWHIAEGCF
CLPWKKILIFKRDQSNERNMSTPIQDNVDQYSELCYTLINRVLCTRPNGS
AEVEYVPCAKERPRESLGGTETYSLLHMPSTDPHARSPEDEYELLMHPRISSHF
LQQRPLMAPSETQFSHL"
CDS
Alignment Scores:
Pred. No.: 4,16e-08 Length: 1654
Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
DB: 9 Gaps: 5
US-10-005-907-2 (1-135) x AF521911 (1-1654)
QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db 173 ATGGGAAATTCCTGCTGAGAGTACTATCTCCATCCAGAGATGCCTTGG 232
QY 20 -----LVLSyGlyAsnProAspGluGluArgLysArg----- 30
Db 233 AATGTGAGATGCAAGACCCCAACAGAGACATCCAGATCTGGATCCACCATATCGCT 292
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 293 GAAGGTGTTCTGCTCCATCGATGAAATAATCTATCTTTTGAAGAGAG---CAAGAT 349
QY 43 GlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnLys 62
Db 350 TCCAAAACGAAATGAAGATGATCTATCTCTCCATCCAGAGCAATGTTGACAGACC 409
QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81
Db 410 TACTCAGAGAGGTGTGCTATATACCTCATCAATCATCGGTCTCTGTACAGGCCATCA 469
QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
Db 470 GGGAACTCTGCTGAAGAGTACTATGAGATGTTCCCTGCAAGAGCTGAGAGACCCAGAG 529
QY 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
Db 530 TCCTTGGGAGAACTGAGACTGAGTATCTACTTCTCATATGCTTCTACAGCCCGAG 589
QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 590 CATGCCCGATCCCAAGAGATGAATGAACTTCTCATGCTCATC 634
RESULT 13
AY212246
LOCUS
DEFINITION Homo sapiens germinal center B-cell expressed transcript 2 (GCET2)
mRNA, complete cds.
ACCESSION AY212246
VERSION AY212246.1 GI:27948576
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE
PubMed

REFERENCE
AUTHORS

TITLE
JOURNAL

FEATURES
source

gene

CDS

ORIGIN

Alignment Scores:
Pred. No.: 8,77e-08 Length: 3270
Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
DB: 9 Gaps: 5

US-10-005-907-2 (1-135) x AY212246 (1-3270)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db 107 ATGGGAAATTCCTGCTGAGAGAAACAGCGCGCAGCAGAACTCAAGAGATGCCTTGG 166
QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db 167 AACGTGAGATGCAAGACCCCAACAGAGAACATCCAGATCTGGATCCCATATCGCT 226
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 227 GAAGGTGTTCTGCTCCATCGATGAAATAATCTATCTTTTGAAGAGAG---CAAGAT 283
QY 43 GlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnLysLys 62
Db 284 TCCAAAACGAAATGAAGATGATCTATCTCTCCATCCAGAGCAATGTTGACAGACC 343
QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81
Db 344 TACTCAGAGAGGTGTGCTATACCTCATCAATCATCGGTCTCTGTACAGGCCATCA 403
QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 3270)
Pan, Z., Shen, Y., Du, C., Zhou, G., Rosenwald, A., Staudt, L. M.,
Greiner, T. C., McKeithan, T. W. and Chan, W. C.
Two newly characterized germinal center B-cell-associated genes,
GCET1 and GCET2, have differential expression in normal and
neoplastic B cells
Am. J. Pathol. 163 (1), 135-144 (2003)

22702315
12819018
2 (bases 1 to 3270)
Pan, Z., Shen, Y., Du, C., Zhou, G., McKeithan, T., Rosenwald, A.,
Staudt, L. and Chan, W. C.
Direct Submission
Submitted (07-JAN-2003) Pathology, University of Nebraska Medical
Center, #42, Omaha, NE 68198, USA
Location/Qualifiers
1..3270
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3q13.13"
/cell_line="DHL16"
1..3270
/gene="GCET2"
/note="GCAT2"
107..643
/gene="GCET2"
/note="similar to Mus musculus germinal center transcript
M17; similar to MGC40441 sequence in GenBank Accession
Number BC030506"
/codon_start=1
/product="germinal center B-cell expressed transcript 2"
/protein_id="AA022147.1"
/db_xref="GI:27948577"
/translation="MGNSLLRNRQNTQEMPNVNRMSQKPTSRCDWHIAEGCF
CLPWKKILIFKRDQSNERNMSTPIQDNVDQYSELCYTLINRVLCTRPNGS
AEVEYVPCAKERPRESLGGTETYSLLHMPSTDPHARSPEDEYELLMHPRISSHF
LQQRPLMAPSETQFSHL"

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Male, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Paney, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
2 (bases 1 to 1532)
Strausberg, R.
Direct Submission
Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapb-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayèle, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W., Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 64 Row: a Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22749536.

FEATURES

Location/Qualifiers
1..1532
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:40441 IMAGE:4385178"
/tissue_type="Lymph, Lymphoma"
/clone_lib="NIH MGC 85"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..1532
/gene="GCET2"
/note="synonyms: MGC40441, HGAL, GCAT2"
/db_xref="LocusID:257144"
/db_xref="MIM:607792"
54..590
/genes="GCET2"
/codon_start=1
/product="germinal center expressed transcript 2"
/protein_id="AAH30506.1"

gene

CDS

Alignment Scores:
Pred. No.: 3.82e-08 Length: 1532
Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
DB: 9 Gaps: 5

US-10-005-907-2 (1-135) x BC030506 (1-1532)

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db 54 ATGGGAAATTCCTGCTGAGAGAAACACAGGGCGGAGAGAACTCAAGATGCTTGG 113
Qy 20 -----LysLysGlyAsnProAspGluArgLysArg----- 30
Db 114 AATGTGAGAAATGCAAGCCCAACAGAGAACTCCAGATGCTGGGATCACCATATCGCT 173
Qy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 174 GAAGGGTGTTCTGCTTCCATCGGAAAAAATACTATTTTGAAGAGG---CAAGAT 230
Qy 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlySer 62
Db 231 TCCCAAAACGAAATGAAGAATGTCATCTACTCCCATCCAGACATGTTGACAGACC 290
Qy 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81
Db 291 TACTCAGAGGAGCTGTGCTATACCTCATCAATCATCGGGTCTCTGTACAGGCCATCA 350
Qy 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
Db 351 GGGAACTCTCTCAAGAGTACTATGAGATGTTCCCTGCAAACTGAGAGACCCAGAG 410
Qy 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
Db 411 TCCTTGGGAGGAACTGAGACTGAGTATTCATCTCTACATATGCTCTTACAGACCCAGG 470
Qy 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 471 CATGCCCGATCCAGAGATGAATATGAATCTTCTCATGCCTCAC 515

RESULT 12
AF521911
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AF521911 1654 bp mRNA linear PRI 14-JAN-2003
Homo sapiens HGAL mRNA, complete cds.
AF521911 GI:27733682
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Lossos, I.S., Alizadeh, A.A., Rajapaksa, R., Tibshirani, R. and Levy, R.
HGAL is a novel interleukin-4-inducible gene that strongly predicts survival in diffuse large B-cell lymphoma
Blood 101 (2), 433-440 (2003)
22397484
PUBMED
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 1654)
Lossos, I.S., Alizadeh, A.A. and Levy, R.
AUTHORS
Direct Submission
TITLE
JOURNAL
Submitted (16-JUN-2002) Division of Oncology, Department of

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

```

/misc_feature /note="assembly_name:Contig2"
64946. .67606
/note="assembly_name:Contig3"
67707. .77674
/note="assembly_name:Contig4"
77775. .89256
/note="assembly_name:Contig5"
89357. .100400
/note="assembly_name:Contig6"
100501. .111915
/note="assembly_name:Contig7"
112016. .125713
/note="assembly_name:Contig8"
125814. .141268
/note="assembly_name:Contig9
clone_end:SP6
vector_side:left"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2.44e-33 Length: 141268
Score: 430.00 Matches: 80
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.14% Indels: 0
DB: 2 Gaps: 0

```

US-10-005-907-2 (1-135) x AC074365 (1-141268)

```

QY 56 GlnGluAsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleAsnHisIle 75
DB 5700 CAGGAAACGAGAAATGGCAGTGTCTTGAAGAAGTGTCTACACTGCTATTATACATC 5641

QY 76 ProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeu 95
DB 5640 CCCCATCAGAGATCCCTCCCTGAGTCCATGATGATGGCTATGAGAACATTCACCTCCCTC 5581

QY 96 ThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThr 115
DB 5580 ACAAGGAAAGTGTAGACAGATTTAGAGAAAGTGTAGACAGACAAATATGCCCTTCTTAGACT 5521

QY 116 SerValSerArgProCysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 5520 TCTGTTAGTAGGCTTGTCTCTCCATCCATGAGCATGATTAAGAGTTGTGTTTCCACAC 5461

```

RESULT 10

```

AL606804 185467 bp DNA linear PRI 23-OCT-2002
LOCUS Human DNA sequence from clone RP11-978115 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL606804
VERSION AL606804.11 GI:24366459
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 185467)

```

REFERENCE

```

AUTHORS Almeida, J.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk
COMMENT During sequence assembly data is compared from overlapping clones.
On Oct 24, 2002 this sequence version replaced gi:18121511.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=

```

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1> RP11-978115 is from the library RPCI-11.4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES

```

source
1..185467
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-978115"
/clone_lib="RPCI-11.4"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 3.29e-33 Length: 185467
Score: 430.00 Matches: 80
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.14% Indels: 0
DB: 9 Gaps: 0

```

US-10-005-907-2 (1-135) x AL606804 (1-185467)

```

QY 56 GlnGluAsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleAsnHisIle 75
DB 132080 CAGGAAACGAGAAATGGCAGTGTCTTGAAGAAGTGTCTACACTGCTATTATACATC 132139

QY 76 ProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeu 95
DB 132140 CCCCATCAGAGATCCCTCCCTGAGTCCATGATGATGGCTATGAGAACATTCACCTC 132199

QY 96 ThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThr 115
DB 132200 ACAAGGAAAGTGTAGACAGATTTAGAGAAAGTGTAGACAGACAAATATGCCCTTCTTAGACT 132259

QY 116 SerValSerArgProCysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 132260 TCTGTTAGTAGGCTTGTCTCTCCATCCATGAGCATGATTAAGAGTTGTGTTTCCACAC 132319

```

RESULT 11

```

BC030506 1532 bp mRNA linear PRI 29-JUN-2004
LOCUS BC030506
DEFINITION Homo sapiens germinal center expressed transcript 2, mRNA (cdna
Clone MGC:40441 IMAGE:4385178), complete cds.
ACCESSION BC030506
VERSION BC030506.1 GI:20987805
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1532)
REFERENCE Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
AUTHORS

```

Research Association for Biotechnology (RAB): cDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers
1..4136
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRACE3002390"
/issue_type="cerebellum"
/clone_lib="BRACE3"
/note="cloning vector: pME18SFL3"
136..537
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC85873.1"
/db_xref="GI:34530325"
/translation="MWLEDSNSTGCGARNVCVCLVFLSVLGCQSRKQGMRTQQAG
RWLPAGREASSETPEGLTLDPSFLAARGWQHEPVRVNCVLPHPFSSWCFCQ
NFKWSETQLNRVTVKRWELIISGSRK"

ORIGIN

Alignment Scores:
Pred. No.: 7.16e-51 Length: 4136
Score: 581.00 Matches: 114
Percent Similarity: 85.19% Conservative: 1
Best Local Similarity: 84.44% Mismatches: 0
Query Match: 81.26% Indels: 20
DB: 9 Gaps: 1

US-10-005-907-2 (1-135) x AK124520 (1-4136)

QY 1 MetGlyAsnTyrLeuLeuArgLysSerCysLeuGlyGluAsnGlnLysLysProLys 20
Db 498 ATGGGAAATTATCTCTCGGAAATCT----- 524
QY 21 LysGlyAsnProaspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
Db 525 -----AGGCAGGAAATGACTACATTTGAAAGAAACTT 557
QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
Db 558 CAAGATCAAGATGAAGAAAGCCAGAGATTCTCCACTTCTATATCAGGAAACGAGAA 617
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db 618 GGCAGTGGTTCTGAAGAAGTGTGTACACTGTCTAATACATCCCCCATCAGAAATCC 677
QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
Db 678 TCCTGAGCTCCAAATGATGCTGTAGAGAACTTGAGAACTTCCCTCACAAGGAAAGTGA 737
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
Db 738 CAGTTTGAAGAAAGGTCAGAGACAGAAATATGCCCTTCTTAGGACTTCTGTTAGGCCT 797
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
Db 798 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCACAC 842

RESULT 9

AC074365/c
LOCUS AC074365 141268 bp DNA linear HTG 23-SEP-2000
DEFINITION Homo sapiens chromosome 1 clone RP11-115C4, WORKING DRAFT SEQUENCE,
10 unordered pieces.
AC074365
ACCESSION AC074365
VERSION AC074365.5 GI:10280935
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

AUTHORS
TITLE
JOURNAL
REFERENCE

1 (Bases 1 to 141268)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (Bases 1 to 141268)
Waterston,R.H.
Direct Submission
Submitted (29-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

COMMENT

On Sep 23, 2000 this sequence version replaced gi:9838075.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0115C04
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-terminator; plasmid; 0%
Chemistry: Dye-terminator ET; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 136274 bases at least Q40
Consensus quality: 138356 bases at least Q30
Consensus quality: 139182 bases at least Q20
Insert size: 142000; agarose-fp
Insert size: 140368; sum-of-contigs
Quality coverage: 4.64 in Q20 bases; agarose-fp
Quality coverage: 4.81 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved

* 1 31377: contig of 31377 bp in length
* 31378 31477: gap of unknown length
* 31478 63011: contig of 31534 bp in length
* 63012 63111: gap of unknown length
* 63112 64845: contig of 1734 bp in length
* 64846 64945: gap of unknown length
* 64946 67606: contig of 2661 bp in length
* 67607 67707: gap of unknown length
* 67707 77674: contig of 9968 bp in length
* 77675 89256: contig of 11482 bp in length
* 89257 89357: gap of unknown length
* 89357 100400: contig of 11044 bp in length
* 100401 100500: gap of unknown length
* 100501 111915: contig of 11415 bp in length
* 111916 112015: gap of unknown length
* 112016 125713: contig of 13698 bp in length
* 125714 125813: gap of unknown length
* 125814 141268: contig of 15455 bp in length.

FEATURES

source
1..141268
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-115C4"
1..31377
/note="assembly_name:Contig10"
31478..63011
/note="assembly_name:Contig11"
63112..64845
misc_feature
misc_feature
misc_feature

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 141268)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (Bases 1 to 141268)
Waterston,R.H.
Direct Submission
Submitted (29-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: Reverse Proteomics Research Institute, HRI and RAB.

FEATURES

source
Location/Qualifiers
1. 4032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRACE3024537"
/tissue_type="cerebellum"
/clone_lib="BRACE3"
/note="Cloning vector: pME18SFL3"

ORIGIN

Alignment Scores:
Pred. No.: 3, 37e-51 Length: 4032
Score: 584.00 Matches: 115
Percent Similarity: 85.19% Conservatives: 0
Best Local Similarity: 85.19% Mismatches: 0
Query Match: 81.68% Indels: 20
DB: 9 Gaps: 1

US-10-005-907-2 (1-135) x AK126682 (1-4032)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
DB 377 ATGGGAATTAATCTCTCTCGAAATCTC----- 403
QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 404 -----AGGCAGGAATGACTACATTTGAAAGAAACTT 436
QY 41 GlnAspGlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
DB 437 CAAGATCAAGATAAGAAAGCCAGAGTTTCTATCCACTTCTAATCAGGAAACGAGAAT 496
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 497 GGCAGTGGTTCTGAAGAAGTGTCTACACTGTCTAATTAATCAGATCCCATCAGATCC 556
QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
DB 557 TCCTGAGCTCCATGATGATGCTATGAGAACATTGACTCCCTCACAGGAAAGTGAGA 616
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
DB 617 CAGTTTAGAGAAAGGTCAGAGACAGAAATATGCCCTTCTTAGGACTTCTGTTAGAGCCT 676
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 677 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTTCCACAC 721

RESULT 7

CQ841482
LOCUS CQ841482 4136 bp DNA linear PAT 02-AUG-2004
DEFINITION Sequence 129 from Patent EP1440981.
ACCESSION CQ841482
VERSION CQ841482.1 GI:50893269
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Nagai, K. and Irie, R.
TITLE Full-length human cdna
JOURNAL Patent: EP 1440981-A 129 28-JUL-2004;
Research Association for Biotechnology (JP)
FEATURES Location/Qualifiers
source 1. 4136
/organism="Homo sapiens"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 7, 16e-51 Length: 4136
Score: 581.00 Matches: 114
Percent Similarity: 85.19% Conservatives: 1
Best Local Similarity: 84.44% Mismatches: 0
Query Match: 81.26% Indels: 20
DB: 6 Gaps: 1

US-10-005-907-2 (1-135) x CQ841482 (1-4136)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
DB 498 ATGGGAATTAATCTCTCTCGAAATCTC----- 524
QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 525 -----AGGCAGGAATGACTACATTTGAAAGAAACTT 557
QY 41 GlnAspGlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
DB 558 CAAGATCAAGATAAGAAAGCCAGAGTTTCTATCCACTTCTAATCAGGAAACGAGAAT 617
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 618 GGCAGTGGTTCTGAAGAAGTGTCTACACTGTCTAATTAATCAGATCCCATCAGAAATCC 677
QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
DB 678 TCCTGAGCTCCATGATGATGCTATGAGAACATTGACTCCCTCACAGGAAAGTGAGA 737
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
DB 738 CAGTTTAGAGAAAGGTCAGAGACAGAAATATGCCCTTCTTAGGACTTCTGTTAGAGCCT 797
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 798 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTTCCACAC 842

RESULT 8

AK124520
LOCUS AK124520 4136 bp mRNA linear PRI 09-SEP-2003
DEFINITION Homo sapiens cDNA FLJ42529 fis, clone BRACE3002390.
ACCESSION AK124520
VERSION AK124520.1 GI:34530324
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

AUTHORS Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, S., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4136)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
COMMENT (E-mail: genomics@hri.co.jp, Tel. 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing:

```

Db      299 TCCCTGAGCTCAATGATGATGGCTATGAGAACATTGACTCCCTCACAGGAAGTGAGA 358
Qy      101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
Db      359 CAGTTTAGAGAAAGGTTCAGAGACAGAAATATGCCCTTCTTAGACTTCTGTAGTAGGCT 418
Qy      121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
Db      419 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCACAC 463

RESULT 5
AB060908      1924 bp      mRNA      linear      PRI 19-MAY-2004
LOCUS      Macaca fascicularis brain cDNA clone:QtrA-14007, full insert
DEFINITION
ACCESSION      AB060908
VERSION      AB060908.1 GI:13874585
KEYWORDS      oligo capping; fis (full insert sequence).
SOURCE      Macaca fascicularis (crab-eating macaque)
ORGANISM      Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.

REFERENCE
1
Osada,N., Hida,M., Kusuda,J., Tanuma,R., Hirata,M., Hirai,M.,
Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
Prediction of unidentified human genes on the basis of sequence
similarity to novel cDNAs from cynomolgus monkey brain
(er) Genome Biol. 3, research0006.1-0006.5 (2001)
2 (bases 1 to 1924)
Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
Direct Submission
Submitted (27-APR-2001) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
(E-mail:khashienih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
Lab host:      TOP10
Vector:      pME18S-FL3 (Acc.No. AB009864)
R. Site:      DraIII (CACTGTGTG)
R. Site:      DraIII (CACCATGTG)
Description: 1st strand cDNA was primed with an oligo(dT) primer
(using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Libraries
were constructed by oligo-capping method
(Sugano et al., Institute of Medical Science, University of
Tokyo).
Custom primer used for sequencing
(5' end primer [CTTCTGCTCTAAAGCTGCG];
3' end primer [CGACCTCGACCTCGAGCAC]).
Location/Qualifiers
1. .1924
/organism="Macaca fascicularis"
/mol_type="mRNA"
/db_xref="taxon:9541"
/clone="QtrA-14007"
/sex="male"
/tissue type="temporal lobe right"
/clone_lib="macaque brain cDNA library QtrA"
/dev_stage="adult"
248..559
/codon_start=1
/product="hypothetical protein"
/protein_id="BAB46903.1"
/db_xref="GI:13874586"
/translation="MTTFERKLQDDKSKVSSISNQENNGSGSEEVYTVNHIP
HRKSSLSNDDGVENIDSLTRKREFRSEYFALLRTSVSRPYSCTHEHDYEVLP
H"

FEATURES
source
CDs
ORIGIN

```

Alignment Scores:

Pred. No.: 8.44e-57 Length: 1924
 Score: 634.00 Matches: 120
 Percent Similarity: 96.12% Conservative: 4
 Best Local Similarity: 93.02% Mismatches: 5
 Query Match: 88.67% Indels: 0
 DB: 9 Gaps: 0

US-10-005-907-2 (1-135) x AB060908 (1-1924)

```

Qy      7   ArgLysLeuSerCysLeuGluAsnGlnLysLysProLysLysGlyAsnProAspGlu 26
Db      170 AGGACATGGATGGCTGGGAGAGAAATCAAAAGAACCCAGGAAGAAACCCAGATGAG 229
Qy      27   GluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLys 46
Db      230 GAAAGAAAACGGCAGGAATGACTACATTTTGAAGAAAATTTCAAGATCAAGATAAGAA 289
Qy      47   SerGlnGluValSerSerThrSerAsnGlnGluAsnGlnGlySerGlySerGluGlu 66
Db      290 AGCAAGAAGTTTCATCCATTCTAATCAGGAAAACGAGAATGSCAGTGGTTCTGAAGAA 349
Qy      67   ValCysTyrThrValIleAsnHisIleProHisGlnArgSerSerSerSerSerAsp 86
Db      350 GTGTGCTACACTGTCAATTAATCATCCCCCATCGGAGGTCTTCCCTGAGCTCCATGAT 409
Qy      87   AspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSer 106
Db      410 GATGGCTATGAGAACATTGACTCCCTCACAGGAAGTGAGAAATTTAGAAAGGTCA 469
Qy      107 GluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCysSerCysThrHisGlu 126
Db      470 GAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGTAGGCTTATCTTGCACCCATGAG 529
Qy      127 HisAspTyrGluValValPheProHis 135
Db      530 CATGATTACGAGTTGTGTTCACAC 556

```

RESULT 6

AKI26682 4032 bp mRNA linear PRI 19-FEB-2004
 LOCUS Homo sapiens cDNA FLJ44728 fis, clone BRACE3024537.
 DEFINITION AKI26682
 ACCESSION AKI26682
 VERSION AKI26682.1 GI:34533254
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
 Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
 Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
 Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
 Sugiyama,T., Irie,K., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
 Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
 Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
 Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
 Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 4032)
 Isogai,T. and Yamamoto,J.
 Direct Submission
 Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

TITLE

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Qy 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 Db 385 TGTTCTCGACCCATGAGCATGATTATGAAGTTGGTTTCCAC 429

RESULT 3
 CQ842314
 LOCUS CQ842314 1709 bp DNA linear PAT 02-AUG-2004
 DEFINITION Sequence 961 from Patent EP1440981.
 ACCESSION CQ842314
 VERSION CQ842314.1 GI:50894101
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Nagai, K. and Irie, R.
 TITLE Full-length human cDNA
 JOURNAL Patent: EP 1440981-A 961 28-JUL-2004;
 Research Association for Biotechnology (JP)
 FEATURES
 source
 1..1709
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 6,12e-65 Length: 1709
 Score: 711.00 Matches: 134
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.26% Mismatches: 0
 Query Match: 99.44% Indels: 0
 DB: 6 Gaps: 0

US-10-005-907-2 (1-135) x CQ842314 (1-1709)

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
 Db 59 ATGGGAATATTATCTCTCGAAACTCAGTTCCTGGAGAGAAATCAAAAGAACCCCAAG 118
 Qy 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
 Db 119 AAAGGAACCCAGATGAGAAAGAAACCGCAGAGAAATGACTACATTTGAAGAAACCTT 178
 Qy 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
 Db 179 CAAGATCGAGATAAGAAAGCCAGAGAGTTTCACTCCACTTCAATCAGGAAACGAGAAAT 238
 Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 Db 239 GGCAGTGGTTCTGAAGAGTGTCTACACTGTCAATTAATCAGATCCCATCAGAGATCC 298
 Qy 81 SerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 Db 299 TCCTGAGTCCCAATGATGCTATGAGACATTTGACTCCCTCCACAGGAAAGTGA 358
 Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 Db 359 CAGTTTAGAGAAAGGTGAGAGACAGAAATATGCTTCTTAGGACTTCTGTAGTGGCT 418
 Qy 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 Db 419 TGTTCTCGACCCATGAGCATGATTATGAAGTTGTGTTTCCAC 463

RESULT 4
 AK123798
 LOCUS AK123798 1709 bp mRNA linear PRI 09-SEP-2003
 DEFINITION Homo sapiens cDNA FLJ41804 fis, clone NOV2000710.
 ACCESSION AK123798
 VERSION AK123798.1 GI:34529425
 KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T., Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R., Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiyama, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E., Kataoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S., Terashima, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE
 AUTHORS Isogai, T. and Yamamoto, J.
 TITLE Direct Submission
 JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB.

FEATURES
 source
 1..1709
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NOVAR2000710"
 /tissue_type="ovary"
 /clone_lib="NOVAR2"
 /dev_stage="adult"
 /note="cloning vector: pME18SFL3"

ORIGIN
 Alignment Scores:
 Pred. No.: 6,12e-65 Length: 1709
 Score: 711.00 Matches: 134
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.26% Mismatches: 0
 Query Match: 99.44% Indels: 0
 DB: 9 Gaps: 0

US-10-005-907-2 (1-135) x AK123798 (1-1709)

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
 Db 59 ATGGGAATATTATCTCTCGAAACTCAGTTCCTGGAGAGAAATCAAAAGAACCCCAAG 118
 Qy 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
 Db 119 AAAGGAACCCAGATGAGAAAGAAACCGCAGAGAAATGACTACATTTGAAGAAACCTT 178
 Qy 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
 Db 179 CAAGATCGAGATAAGAAAGCCAGAGAGTTTCACTCCACTTCAATCAGGAAACGAGAAAT 238
 Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 Db 239 GGCAGTGGTTCTGAAGAGTGTCTACACTGTCAATTAATCAGATCCCATCAGAGATCC 298
 Qy 81 SerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100

Db 348 CAGTTTATAGAAAGGTCAGACAGATATGCGCTTCTTAGGACTTCTGTTAGTAGGCGCT 407
 Qy 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 Db 408 TGTTCCTGCACCCATGACGATGATTATGAAGTTGTGTGTTCCACAC 452

RESULT 2
 AX505122
 LOCUS AX505122 3762 bp DNA linear PAT 27-SEP-2002
 DEFINITION Sequence 1 from Patent WO0246389.
 ACCESSION AX505122
 VERSION AX505122.1 GI:23386429
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS Nocka, K., Birozzi, G. and Einstein, R.
 TITLE Novel genes associated with allergic hypersensitivity and mast cell
 activation
 JOURNAL Patent: WO 0246389-A 1 13-JUN-2002;
 UCB, S.A. (BE)

FEATURES
 source Location/Qualifiers
 1..3762
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 25..432
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD48808.1"
 /db_xref="GI:23386430"
 /translation="MGNYLLRLKSLGLENQKKPKGNPDPEERKROEMTTFERKIQDDQ
 KKSQEVSTSNQENENGSGEEVCYTVINHI PHORSSLSNDGYNIDSLTRKVFQF
 RERSETEYALLRTSVSRPCSCTHEHDYEVVFEH"

CDS
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,52e-65 Length: 3762
 Score: 715.00 Matches: 135
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-005-907-2 (1-135) x AX505122 (1-3762)

Qy 1 MetGlyAsnTyrIleuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
 Db 25 ATGGGAAATATATCTCTCGAAAACTCAGTTGCTGGGAGAAATCAAAAGAACGCCAAG 84
 Qy 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
 Db 85 AAAGGAAACCCAGATCAGAGAAAGAAAACGACAGGAATGACTACATTTCGAAGAAACATT 144
 Qy 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
 Db 145 CAAGATCAAGATAAAGAAAGCCAAAGTTTCATCCACTTCTTAATCAGGAAAAACGAGAT 204
 Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 Db 205 GGCAGTGGTTCGAAGAAGTGGCTACACTGTCATTAAATCATCATCCCATCAGATCC 264
 Qy 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 Db 265 TCCTTGAGCTCCAAATGATGCGCTATGAGAAACATGACTCCCTCACAAGGAAGTGAGA 324
 Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 Db 325 CAGTTTATAGAAAGGTCAGACAGATATGCGCTTCTTAGGACTTCTGTTAGTAGGCGCT 384

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 18, 2004, 21:31:29 ; Search time 3546 Seconds
(without alignments)
1800.369 Million cell updates/sec

Title: US-10-005-907-2
Perfect score: 715
Sequence: 1 MGNVLLRLKSLGKQKPK.....SVSRPCSTHEHDYEVVFPF 135

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-O=/cgn2 1/USPTO spool/US10005907/runat 16112004 140457 6169/app query.fasta 1.327
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10005907 @CGN 1 1 3731 @runat 16112004 140457 6169 -NCPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	715	100.0	2648	9 BC024174	BC024174 Homo sapi
2	715	100.0	3762	6 AX505122	AX505122 Sequence
3	711	99.4	1709	6 CQ842314	CQ842314 Sequence
4	711	99.4	1709	9 AK123798	AK123798 Homo sapi

ALIGNMENTS

RESULT 1
BC024174
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BC024174
Homo sapiens LOC148823, mRNA (cdna clone IMAGE:4109064).
BC024174
BC024174.1 GI:18848218

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 2648)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stratton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullahy,S.J., Bosak,S.A., McSwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

AB060908 Macaca fa
AK126682 Homo sapi
CQ841482 Sequence
AK124520 Homo sapi
AC074355 Homo sapi
AL060804 Human DNA
BC030506 Homo sapi
AF521911 Homo sapi
AY212246 Homo sapi
CQ733169 Sequence
AX740457 Sequence
AY500830 Mus muscu
U13263 Mus muscu
AK125521 Homo sapi
AK128688 Homo sapi
AC024964 Homo sapi
AC110870 Homo sapi
AL034556 Plasmid
AC108574 Rattus no
AC104910 Mus muscu
AC124636 Mus muscu
AC139696 Homo sapi
AL954353 Mouse DNA
AL147571 Pan trogl
AL626775 Mouse DNA
AL713872 Mus muscu
AC068147 Homo sapi
AC005801 Leishmani
AC009603 Leishmani
AC121231 Leishmani
BD136402 95 human
AC092331 Homo sapi
AC116364 Homo sapi
AC069163 Homo sapi
AC004709 Plasmid
AE014847 Plasmid
AC120746 Rattus no
AC108226 Rattus no
AL837509 Mouse DNA
AC110159 Mus muscu

Search completed: November 19, 2004, 03:36:44
Job time : 2710 secs

```

Alignment Scores:
  Pred, No.:      5,51e-08      Length:      508
  Score:         156,50        Matches:      37
  Percent Similarity: 60.75%    Conservative: 28
  Best Local Similarity: 34.58% Mismatched:   39
  Query Match:    21.89%       Indels:       3
  DB:             9           Gaps:         3

US-10-005-907-2 (1-135) x AY404586 (1-508)

QY      31  GlnGluMetThrThrPheGluArgLysLeuGlnAspLysLysSerGlnGluVal 50
DB      116  AAAAAAATACTCAATTTTTGAAAGAGG---CAAGATTCCTCCAAACGAAATGAAGAATG 172
QY      51  SerSerThrSerAsnGlnGluAsnGluAsnGlySerGlySerGluGluValCysrThr 70
DB      173  TCATCTACTCCCATCCAGGACAAATGTTGACACAGACCTACTCAGAGGAGCTGTGCTATACC 232
QY      71  ValIleAsnHis---IleProHisGlnArgSerLeuSerSerAsnAspAspGlyTyr 89
DB      233  CTATCAATCATCGGGTCTCTGTACAGGCCATCAGGGAACTCTGCTGAAGAGTACTAT 292
QY      90  GluAsnTleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu 109
DB      293  GAGAAATGTTCCCTGCMAAGCTGAGAGAGCCAGAGAGTCTCTTGGGAGGAACTGAGACTGAG 352
QY     110  TyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp 128

```

Qy	129	TyrGluValValPheProHis 135 ::: :::
Dd	413	TATGAACCTTCATGCCCTCAC 433
RESULT 14		
AW9633366		
LOCUS	AW9633366	683 bp mRNA linear EST 01-JUN-2000
DEFINITION	ESR3754339 MAGE resequences, MAGH Homo sapiens CDNA, mRNA sequence.	
ACCESSION	AW9633366	
VERSION	AW9633366.1	GI:8153202
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 683) Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J. Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray	
AUTHORS	Unpublished (2000)	
TITLE	Contact: John Quackenbush	
JOURNAL	The Institute for Genomic Research	
COMMENT	9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 3528	

```

FEATURES
  source
    1..883
      Location/Qualifiers
        organism="Homo sapiens"
        mol_type="mRNA"
        db_xref="taxon:9606"
        clone_lib="WAGE resequences, MAGH"
        note="Vector: pBluescriptSKm"

ORIGIN

Alignment Scores:      8.08e-08      Length:      683
Pred. No.:             156.50      Matches:    37
Score:

```

```

QY 19 Pro-----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db 125 CTTTGAATGTCAGATGCAAGGCCCAACAGAGAACATCCAGATGCTGGGATCACCAT 184
QY 31 -----GlnGluMetThrThrPheGluArgLysLeu 40
Db 185 ATCGCTGAAGGCTGTTCTGCTCCATGTAAGAAATACTCATTTTGAAGAGG--- 241
QY 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
Db 242 CAAGATCCCAACAGAAATGAAAGATGTCATCTACTCCATCCAGGCAATGTCAC 301
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArg 79
Db 302 CAGACCTACTCAGAGGAGCTGTCTATACCTCATCATCTCGGTTCTCTGTACAAAG 361
QY 80 SerSerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal 99
Db 362 CCATCAGGGAATCTGCTGAAGAGTACTATGAGAATGTTCCCTGCAAGCTGAGAGACC 421
QY 100 ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
Db 422 AGAGAGCTCTGGAGAGAGTACTGAGTATGATTCATCTACATATGCTTCTACAGAC 481
QY 120 ProCysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 482 CCCAGGATGCCGATCCCAAGAGATGAATATGAACTTCTCATGCTCTCAC 532

RESULT 12
BM455407
LOCUS
DEFINITION
AGENCOURT 6407334 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500391
5', mRNA sequence.
BM455407
BM455407.1 GI:18504436
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2135 row: d column: 24
High quality sequence start: 19
High quality sequence stop: 760.
Location/Qualifiers
1. 1085
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5500391"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source
1..508
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>508
/gene="GCET2"
/locus_tag="HCM1940"

ORIGIN
Alignment Scores:

```

```

Pred. No.: 6.55e-08 Length: 1085
Score: 159.50 Matches: 37
Percent Similarity: 60.75% Conservative: 28
Best Local Similarity: 34.58% Mismatches: 39
Query Match: 22.31% Indels: 3
DB: 4 Gaps: 3

US-10-005-907-2 (1-135) x BM455407 (1-1085)
QY 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGluVal 50
Db 439 AAAAAAATCTCATTTTGAAGAGG---CAAGATTCCCAACGAAATGAAGATG 495
QY 51 SerSerThrSerAsnGlnGluAsnGlySerGlySerGluGluValCysTyrThr 70
Db 496 TCATCTACTCCATCCAGGCAATGTTGACCAAGCTACTCAGAGGAGCTGTGTATACC 555
QY 71 ValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspGlyTyr 89
Db 556 CTCATCAATCATCGGTTCTCTGTACAAAGCCATCANGAACTCTGCTGAAGAGTACTAT 615
QY 90 GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu 109
Db 616 GAGATGTTCTCCGCAAGCTGAGAGACCCAGAGAGTCTTGGGAGAGAACTGAGACTGAG 675
QY 110 TyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp 128
Db 676 TATTCACCTTCTACATATGCTTCTACAGACCCAGGATGCCGATCCCAAGAGATGAA 735

RESULT 13
AY404586
LOCUS
DEFINITION
Homo sapiens GCET2 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY404586
AY404586.1 GI:39760563
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 508)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
Perriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 508)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
Perriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..508
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>508
/gene="GCET2"
/locus_tag="HCM1940"

ORIGIN

```

Qy 61 GlySerGlySerGlucIuValCysTyrThrValIIleAsnHis---lleProHISGLNArg 79
Db 373 CAGACCTACTCAGAGAGCTGTCTATACCCTCATCAATCATCGGGTTCTCTACAAGG 433

Qy 80 SerSerLeuSerSerAsnAaspGlyTyrGIUasnlEaspSerLeuThraRgLYseVal 99
Db 433 CCATAGGAACCTCTCTCAAGAGTAGTATGAGAATGTTCCCTGCAAAGCTGAGAGACCC 492

Qy 100 ArgGINPheArGluArGserGluThrGluTyRAlaleuLeuargThrSerValSerArg 119
Db 493 AGAGAGTCCTTGGAGAACACTGAGACTGAGTAGTATTCTACTTCTACATATGCCCTTCACAGAC 552

Qy 120 ProCysSerCys---ThrHiSGLIuHiASpTYRGluValValPheProHiS 135
Db 553 CCCAGGCATGCCCGATCCCCAGAGAGTAGATATGAACCTTCATGCCCTCAC 603

```

RESULT 11
BM455198                1193 bp      mRNA       linear    EST 05-FEB-2002
AGENCOURT 6405645 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500141
5'', mRNA sequence.
BM455198
BM455198.1 GI:18504238
EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1193)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llni.gov
Plate: LHAM12134 row: j column: 14
High quality sequence stop: 592.
Location/Qualifiers
location=1..1193
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies."
Note: this is a NIH_MGC Library."
```

ORIGIN

Alignment Scores:		
Pred. No.:	1.12e-08	Length: 1193
Score:	166.50	Matches: 48
Percent Similarity:	52.23%	Conservative: 34
Best Local Similarity:	30.57%	Mismatches: 52
Query Match:	23.29%	Indels: 23
DB:	4	Gaps: 6

US-10-005-907-2 (1-135) x BM455198 (1-1193)

Qy 1 MetGlyAsnTyrLeuLeuArgylsLeuSerCysLeuGlyGlu-----AsnGlnLysLys 18
Db 65 ATGGCAAATTCCTCTCTGTCAGAGAAAACAGTTTTTCAGGCGGAGCAGCAACACTCAAGAGATG 124

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1566)

REFERENCE

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

COMMENT

FEATURES

Location/Qualifiers
1..1566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODG007YD15"
/tissue type="B cells (Ramos cell line)"
/plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:
Pred. No.: 7.12e-09 Length: 1566
Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
DB: 3 Gaps: 5

US-10-005-907-2 (1-135) x CR625820 (1-1566)

QY 1 MetGlyAsnTyrLeuArgLysLeuSerCysLeuGluAsnGlnLysLysPro--- 19
Db 115 ATGGGAATTCCTGCTGAGAGAAACAGCGCGCAGCAGAACACTCAAGAGATGCTTGG 174
QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db 175 AATGTGAGATGCAAGACCCCAACAGACAGACATCCAGATGCTGGGATCCATATCGCT 234
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 235 GAAGGGTGTTCGCTTCCATCGAAAAAATACTCATTTTGAAGAAGG---CAAGAT 291
QY 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnGlySer 62
Db 292 TCCCAACGAAATGAAAGATGTCTACTCTCCATCCATCCAGACAAATGTTGACCGACC 351
QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81
Db 352 TACTCAGAGAGCTGTGTATACCTCATCAATCATCGGTTCTCTGTACAGCCATCA 411
QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
Db 412 GGAAGTCTGCTGAGAGTACTATGAGAAATGTTCCCTGCAAGCTCAGAGACCCAGAG 471
QY 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
Db 472 TCCTTGGGAGGAACTGAGACTGAGTATGATTCATCTTACATATGCTTCTACAGACCC 531
QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 532 CATGCCCGATCCCAAGAGATGAATATGAATTCATCTTCATGCCCTCAC 576

RESULT 9

BM904106 1000 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT.6692745 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5499556
DEFINITION 5', mRNA sequence.
ACCESSION BM904106
VERSION BM904106.1 GI:19354351
KEYWORDS EST.

SOURCE

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1000)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12133 row: b column: 05
High quality sequence stop: 698.
Location/Qualifiers
1..1000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5499556"
/tissue type="lymphoma, cell line"
/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC_85"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Alignment Scores:
Pred. No.: 8.95e-09 Length: 1000
Score: 165.50 Matches: 48
Percent Similarity: 52.23% Conservative: 34
Best Local Similarity: 30.57% Mismatches: 52
Query Match: 23.29% Indels: 23
DB: 4 Gaps: 6

US-10-005-907-2 (1-135) x BM904106 (1-1000)

QY 1 MetGlyAsnTyrLeuArgLysLeuSerCysLeuGlyGlu-----AsnGlnLysLys 18
Db 111 ATGGGAATTCCTGCTGAGAGAAACAGTTTCAGCGCGCAGCAGAACACTCAAGAGATG 170
QY 19 Pro-----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db 171 CCTTGAATGTGAGATGCAAGACCCCAACAGACACATCCAGATGCTGGATCACCAT 230
QY 31 -----GlnGluMetThrThrPheGluArgLysLeu 40
Db 231 ATCGCTGAAGGGTGTTCGCTTCCATCGAAAAAATACTCATTTTGAAGAAGG--- 287
QY 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGln 60
Db 288 CAAGATTCCTCCAAACGAAATGAAAGATGCTATCTACTCCATCCATCCAGACAATGTGAC 347
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArg 79
Db 348 CAGACCTACTCAGAGGAGCTGTGTATACCTCATCATCATCATCGGTTCTCTGTACAAG 407
QY 80 SerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal 99
Db 408 CCATCAGGAACTCTGCTGAAGAGTACTATGAGATGTTCCCTGCAAGAGCTGAGAGACC 467
QY 100 ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
Db 468 AGAGAGTCTCTGGGAGGAACTGAGACTGAGTATTCACTTCTACATATGCTTCTACAGAC 527

ORIGIN		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:4342719"	
		/tissue_type="lymphoma, cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_85"	
		/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."	
US-10-005-907-2 (1-135) x BF795688 (1-1124)			
QY	1	MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro	19
Db	43	ATGGGAAATTCCTCTGAGAGAAACACGGCGGACGAGACACTCAAGAGATGCCTTGG	102
QY	20	-----LysLysGlyAsnProAspGluGluArgLysArg	30
Db	103	AATGTGAGAAATGCAAGACCCCAACACAGAGAACATCCAGATGCTGGGATCACCATATCGCT	162
QY	31	-----GlnGluMetThrThrPheGluArgLysLeuGlnAsp	42
Db	163	GAAGGGTGTTCCTGCTCCATGGAATAAATACTATTTTGAAGAAGG---CAAGAT	219
QY	43	GlnAspLysLysSerClnGluValSerThrSerAsnGlnGluAsnGlnGlySer	62
Db	220	TCCCAAAACGAAATGAAAGATGCTATCTATCTCCATCCAGACAAATGTTGACACAGACC	279
QY	63	GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer	81
Db	280	TACTCAGAGAGCTGTGCTATACCTCATCATCATCTCGGTTCTCTGTCAAGGCCATCA	339
QY	82	LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln	101
Db	340	GGGAACCTCTGCTGACAGACTATGAGAAATGTTCCCTGCAAGAGCTGAGAGACCCAGAG	399
QY	102	PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys	121
Db	400	TCCTTGGGAGGAACTGAGACTGAGTATTCATCTTCATATGCTTCTACAGACCCGAG	459
QY	122	SerCys---ThrHisGluHisAspTyrGluValValPheProHis	135
Db	460	CATGCCCGATCCCAAGAGATGAATATGAATCTTCATGCTCTCATGCTCAC	504
RESULT 7			
BF795688		1124 bp mRNA linear EST 12-JAN-2001	
LOCUS		602259560F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342719 5',	
DEFINITION		mRNA sequence.	
ACCESSION		BF795688	
VERSION		BF795688.1 GI:12100659	
KEYWORDS		EST.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		NIH-MGC http://imgc.ncbi.nih.gov/.	
JOURNAL		National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT		Unpublished (1999)	
		Contact: Robert Strauberg, Ph.D.	
		Email: cgapbs-remail.nih.gov	
		Tissue Procurement: Louis Staudt, M.D., Ph.D.	
		cDNA Library Preparation: Life Technologies, Inc.	
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
		DNA Sequencing Arrayed by: Incyte Genomics, Inc.	
		Clone distribution: MGC clone distribution information can be	
		found through the I.M.A.G.E. Consortium/LLNL at:	
		http://image.llnl.gov	
		Plate: L1AM9958 row: h column: 16	
		High quality sequence stop: 715.	
FEATURES		Location/Qualifiers	
source		1..1124	
Alignment Scores:			
Pred. No.:	4.49e-09	Length:	1095
Score:	169.50	Matches:	47
Percent Similarity:	51.61%	Conservative:	33
Best Local Similarity:	30.32%	Mismatches:	54
Query Match:	23.71%	Indels:	21
DB:	4	Gaps:	5
US-10-005-907-2 (1-135) x BG106563 (1-1095)			
QY	1	MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro	19
Db	43	ATGGGAAATTCCTCTGAGAGAAACACGGCGGACGAGACACTCAAGAGATGCCTTGG	102
QY	20	-----LysLysGlyAsnProAspGluGluArgLysArg	30
Db	103	AATGTGAGAAATGCAAGACCCCAACACAGAGAACATCCAGATGCTGGGATCACCATATCGCT	162
QY	31	-----GlnGluMetThrThrPheGluArgLysLeuGlnAsp	42
Db	163	GAAGGGTGTTCCTGCTCCATGGAATAAATACTATTTTGAAGAAGG---CAAGAT	219
QY	43	GlnAspLysLysSerClnGluValSerThrSerAsnGlnGluAsnGlnGlySer	62
Db	220	TCCCAAAACGAAATGAAAGATGCTATCTATCTCCATCCAGACAAATGTTGACACAGACC	279
QY	63	GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer	81
Db	280	TACTCAGAGAGCTGTGCTATACCTCATCATCATCTCGGTTCTCTGTCAAGGCCATCA	339
QY	82	LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln	101
Db	340	GGGAACCTCTGCTGACAGACTATGAGAAATGTTCCCTGCAAGAGCTGAGAGACCCAGAG	399
QY	102	PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys	121
Db	400	TCCTTGGGAGGAACTGAGACTGAGTATTCATCTTCATATGCTTCTACAGACCCGAG	459
QY	122	SerCys---ThrHisGluHisAspTyrGluValValPheProHis	135
Db	460	CATGCCCGATCCCAAGAGATGAATATGAATCTTCATGCTCTCATGCTCAC	504
RESULT 8			
BF795688		1566 bp mRNA linear HTC 21-JUL-2004	
LOCUS		602259560F1 NIH_MGC_85 Homo sapiens cDNA clone CSODG007YD15 of B cells (Ramos cell line) of	
DEFINITION		full-length cDNA clone CSODG007YD15 of B cells (Ramos cell line) of	
ACCESSION		BF795688	
VERSION		BF795688.1 GI:50506627	
KEYWORDS		HTC; CNSLT; cDNA.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		1 (bases 1 to 1566)	
JOURNAL		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
COMMENT		Full-length cDNA libraries and normalization	
		Unpublished	
		Contact : Feng Liang Email : fliang@lifetech.com URL :	

```

Qy 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlySer 62
Db 273 TCCAAAACGAAATGAAGAATGTCATCTACTCCATCCAGCAATGTTGACAGACC 332
Qy 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81
Db 333 TACTCAGAGGAGCTGTGTATACCTCATCAATCATCGGTTCTCTGTACAGGCCATCA 392
Qy 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
Db 393 GGGAACTCTGCTGAAGACTACTCAGAAATGTTCCCTGCAAGCTGAGAGACCCAGAG 452
Qy 102 PheArgGluArgSerGluThrGluValAlaLeuValPheProHis 135
Db 453 TCCTTGGGAGGAAGTCTGAGCTGAGTATTCATCTTACATATGCTTCTACAGACCCAGG 512
Qy 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 513 CATGCCGATCCCGAAGATGAATATGAATCTCTCATGCCTCAC 557

RESULT 5
BM479887
LOCUS
DEFINITION
AGENCOURT_6464830 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5577364
5', mRNA sequence.
BM479887
VERSION
BM479887.1 GI:18528929
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
found through the MGC clone distribution information can be
http://image.llnl.gov
Plate: LAM12330 row: 1 column: 05
High quality sequence stop: 665.
FEATURES
Location/Qualifiers
1..1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5577364"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.: 4 19e-09 Length: 1039
Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
DB: 4 Gaps: 5
US-10-005-907-2 (1-135) x BM479887 (1-1039)

```

```

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db 81 ATGGGAAATTCCTGCTCAGAGAAAACAGGGCGGAGCAGAAACACTCAAGAGATGCTTGG 140
Qy 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db 141 AATGTGAGATGCAAGACCCCAACAGAGAACATCCAGATGCTGGGATCACCATATCGCT 200
Qy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 201 GAAGGGTGTTCCTGCCCTTCCATGGAAAAAATACTATCTTTTGAAGAAGG---CAGAT 257
Qy 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlySer 62
Db 258 TCCAAAACGAAATGAAGAATGTCATCTACTCCATCCAGGACAAATGTTGACAGACC 317
Qy 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81
Db 318 TACTCAGAGGAGCTGTGTATACCTCATCAATCATCGGTTCTCTGTACAGGCCATCA 377
Qy 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
Db 378 GGGAACTCTGCTGAAGACTACTATGAGAATGTTCCCTGCAAGCTGAGAGACCCAGAG 437
Qy 102 PheArgGluArgSerGluThrGluValAlaLeuValPheProHis 121
Db 438 TCCTTGGGAGGAAGTCTGAGCTGAGTATTCATCTTACATATGCTTCTACAGACCCAGG 497
Qy 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 498 CATGCCGATCCCGAAGATGAATATGAATCTCTCATGCCTCAC 542

BG106563 1095 bp mRNA linear EST 30-JAN-2001
602290373F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4385178 5',
mRNA sequence.
BG106563
GI:12600409
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM10066 row: a column: 19
High quality sequence stop: 641.
FEATURES
Location/Qualifiers
1..1095
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4385178"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 85"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

```

VERSION      AL560376.3  GI:46185747
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 917)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31284507.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1964.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DG007CB080P1&c=1964.r.
FEATURES             source
    1..917
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0DG007YD15"
        /tissue_type="B CELLS (RAMOS CELL LINE)"
        /cell_line="RAMOS CELL LINE"
        /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
        /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
        with a NotI-oligo(dT) primer. Five prime end enriched,
        double-strand cDNA was digested with Not I and cloned into
        the Not I and EcoRV sites of the pCMVSPORT 6 vector.
        Library was not normalized."
ORIGIN
Alignment Scores:      3.57e-09      Length:      917
Pred. No.:            169.50      Matches:      47
Score:                 51.61%      Conservative:  33
Percent Similarity:    30.32%      Mismatches:   54
Best Local Similarity: 23.71%      Indels:       21
Query Match:          1
DB:
US-10-005-907-2 (1-135) x AL560376 (1-917)

QY      1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db      115 ATGGGAATTTCTCTGCTGAGAGAAACAGCGCGCAGAGAACACTCAAGAGATGCCTTG 174
QY      20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db      175 AATGTGAGATGCAAGAGCCCAACAGAGAACATCCAGATGCTGGGATCCACATATCGCT 234
QY      31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db      235 GAAGGGTGTTCCTCCCTTCCATCGAAAAAATACTCATTTTGAAGAGAGG---CAAGAT 291
QY      43 GlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlnLysSer 62
Db      292 TCCCAAAACGAAATGAAGAATGTCTACTCTCCCATCCAGGACAAATGTTGACGAGACC 351
QY      63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81
Db      352 TACTCAGAGAGCTGTGCTATACCTCTCATCATCATCGGGTCTCTGTACAAAGCCCATCA 411
QY      82 LeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
Db      412 GGGAACTCTGCTGAAGAGTACTATGAGAAATGTTCCCTCCAAAGCTGAGAGACCCAGAGAG 471
QY      102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
Db      472 TCCTTGGGAGGAGTACGAGCTGAGTATTCTACATATGCCTTCTACAGACCCAGG 531
QY      122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db      532 CATGCCCGATCCCCAGAGATGAATATGAATCTTCTATGCCTTCAC 576
QY      532 CATGCCCGATCCCCAGAGATGAATATGAATCTTCTATGCCTTCAC 576
Db      532 CATGCCCGATCCCCAGAGATGAATATGAATCTTCTATGCCTTCAC 576
RESULT 4
BG389415
LOCUS     602414106F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4522512 5',
DEFINITION mRNA sequence.
ACCESSION BG389415
VERSION   BG389415.1 GI:13282861
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 924)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LLML0423 row: 1 column: 01
High quality sequence stop: 726.
FEATURES             Location/Qualifiers
    1..924
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4522512"
        /tissue_type="embryonal carcinoma, cell line"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH MGC 92"
        /note="Organ: Testis; Vector: pCMV-SPORT6; Site 1: NotI;
        Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
        Average insert size 2.5 kb. Library enriched for
        full-length clones and constructed by Life Technologies.
        Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:      3.6e-09      Length:      924
Pred. No.:            169.50      Matches:      47
Score:                 51.61%      Conservative:  33
Percent Similarity:    30.32%      Mismatches:   54
Best Local Similarity: 23.71%      Indels:       21
Query Match:          4
DB:
US-10-005-907-2 (1-135) x BG389415 (1-924)

QY      1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db      96 ATGGGAATTTCTCTGCTGAGAGAAACAGCGCGCAGAGAACACTCAAGAGATGCCTTG 155
QY      20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db      156 AATGTGAGATGCAAGAGCCCAACAGAGAACATCCAGATGCTGGGATCCACATATCGCT 215
QY      31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db      216 GAAGGGTGTTCCTCCCTTCCATCGAAAAAATACTCATTTTGAAGAGAGG---CAAGAT 272

```


/tissue type="from acute myelogenous leukemia"
/lab host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_55"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI
(ggcattatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CACGGCCATTAGGCC-3' and
3' adaptor sequence:
5'-ATTCTAGAGCGGAGCGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."

ORIGIN

Alignment Scores:
Pred. No.: 2.78e-73 Length: 527
Score: 715.00 Matches: 135
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-005-907-2 (1-135) x BF242113 (1-527)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAenGlnLysLysProLys 20
DB 45 ATGGGAAATTTCTCTCGGAAATCTAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 104
QY 21 LysGlyAsnProAspGluGluArgLysLeuGlnMetThrThrPheGluArgLysLeu 40
DB 105 AAAGGAAACCCAGATGAGGAAAGAAACGGCAGGAAATGACTATTGGAAGAAACTT 164
QY 41 GlnAspGlnAspLysLeuSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
DB 165 CAAGATCAAGATAAGAAAGCCAGAGAGTTTCATCCACTTCTATATCAGGAAACGAGAT 224
QY 61 GlySerGlySerGluGluValCysTyrThrValleAsnHisLeProHisGlnArgSer 80
DB 225 GGCAGTGGTCTGGAAGAGTGTGTACACTGTCTAATTAATCACATCCCATCCAGATCC 284
QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnLeuAspSerLeuThrArgLysValArg 100
DB 285 TCCCTGAGCTCCATGATGATGCTATGAGAACATTCCTCCCTCAAGGAAAGTGAGA 344
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
DB 345 CAGTTTAGAGAAAGTTCAGAGACAGATAATGCCCTTCTAGGACTTCTGTAGTGGCCT 404
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 405 TGTTCTGCACCCATGAGCATGATTATGAAGTTGTGTGTTCCACAC 449

RESULT 2
BF797507 1045 bp mRNA linear EST 12-JAN-2001
LOCUS 602257174F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4340323 5',
DEFINITION mRNA sequence.
ACCESSION BF797507
VERSION BF797507.1 GI:12102561
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1045)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM9952 row: d column: 20
High quality sequence stop: 687.

FEATURES

source
Location/Qualifiers
1..1045
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4340323"
/tissue_type="lymphoma, cell line"
/lab host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 1.44e-09 Length: 1045
Score: 173.50 Matches: 48
Percent Similarity: 51.61% Conservative: 32
Best Local Similarity: 30.97% Mismatches: 54
Query Match: 24.27% Indels: 21
DB: 2 Gaps: 5

US-10-005-907-2 (1-135) x BF797507 (1-1045)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAenGlnLysLysPro--- 19
DB 103 ATGGGAAATTTCTCTGCTGAGAGAAACAGCGCGCAGAGAACACTCAAGAGATGCCTTGG 162
QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
DB 163 AATGTGAGATGCAAGACCCCAACAGAGAACATCCAGATGCTGGGATCACCATATCGCT 222
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
DB 223 GAAGGGTGTTCCTGCTTCCATCGAAAAAATACTCATTTTTCAGAAAGAGG---CAAGAT 279
QY 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAenGluAenGlySer 62
DB 280 TCCCAAAACAAATAATGAAAGAATGTCTATCTATCCCATCCAGGACAAATGTTGACACAGCC 339
QY 63 GlySerGluGluValCysTyrThrValleAsnHis---IleProHisGlnArgSerSer 81
DB 340 TACTCAGAGGAGCTGTCTATACCTTCATCATTCATCGGGTCTCTGTGTACAGGCCATCA 399
QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
DB 400 GGGAACTCTCTCAAGAGTACTATGAGATGTTCCCTGCAAGAGCTGAGAGACCCAGAG 459
QY 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
DB 460 TCTTGGGAGGAACTGAGAGTATGAGTATTCATCTTCTACATATGCCTTCTACAGACCCAGG 519
QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
DB 520 CATGCCCGATCCCGAGAGATGAATGATGAATTCATGCTTCATGCCTCAC 564

RESULT 3
AL560376
LOCUS
DEFINITION
ACCESSION

AL560376 917 bp mRNA linear EST 02-APR-2004
AL560376 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
clone GS0DG007YD15 5-PRIME, mRNA sequence.
AL560376

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 19, 2004, 01:11:39 ; Search time 2705 Seconds
(without alignments)
1818.619 Million cell updates/sec

Title: US-10-005-907-2
Perfect score: 715
Sequence: 1 MGNVLLRLKSLCNGENQKPKP.....SVSRPCSTHEHDYEVVFP 135

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cpn2.1/USPTO.spool/US10005907/runat.16112004.140457.6176/app.query.fasta.1.327
-DB=EST -QFMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10005907 @CGN 1.1 3437 @runat.16112004.140457.6176 -NCPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic.*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	715	100.0	527	2	BF242113
2	173.5	24.3	1045	2	BF797507
3	169.5	23.7	917	1	AL560376
4	169.5	23.7	924	4	EG389415
5	169.5	23.7	1039	4	EG479887
6	169.5	23.7	1095	4	EG106563
7	169.5	23.7	1124	2	BF795688
8	169.5	23.7	1566	3	CR625820
9	166.5	23.3	1000	4	BN904106

10	166.5	23.3	1070	4	BM456595
11	166.5	23.3	1193	4	BM455198
12	159.5	22.3	1085	4	BM455407
13	156.5	21.9	508	9	AY404586
14	156.5	21.9	683	2	AW963366
15	156.5	21.9	876	5	BQ214406
16	150.5	21.0	1041	4	BQ24891
17	149	20.8	1165	4	BM458339
18	146.5	20.5	1059	4	BM459235
19	145	20.3	495	1	AA311038
20	142	19.9	526	6	CB286394
21	133.5	18.7	522	4	BG149093
22	128	17.9	825	6	CB320849
23	127	17.8	520	4	BG145241
24	127	17.8	524	4	BG276802
25	127	17.8	526	4	BG276546
26	125	17.5	661	2	BB629482
27	125	17.5	3317	3	AK036553
28	122	17.1	560	5	BU932945
29	118.5	16.6	815	5	BU932945
30	118	16.5	828	5	BX390865
31	116.5	16.3	998	4	BG176880
32	112	15.7	479	5	BU429114
33	110.5	15.5	539	2	BG276071
34	109.5	15.3	538	2	AW519547
35	108	15.1	1015	4	BM477520
36	105.5	14.8	1527	9	AG032114
37	104.5	14.6	502	9	AY404588
38	104	14.5	591	2	BF797042
39	103.5	14.5	927	4	BG177483
40	102.5	14.3	833	7	CK397921
41	102.5	14.3	965	5	BX391573
42	101.5	14.2	288	2	AW480906
43	101.5	14.2	875	9	AG181241
44	100.5	14.1	508	9	CE686854
45	100	14.0	594	8	BH085248

ALIGNMENTS

RESULT 1

BF242113

LOCUS

DEFINITION

BF242113

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BF242113 527 bp mRNA linear EST 14-NOV-2000
601880401F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4109064 5',
mRNA sequence.
BF242113 GI:11156040
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 527)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LICM993 row: a column: 01
High quality sequence stop: 518.
Location/Qualifiers
1. 527
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4109064"

CC hyperproliferative disorders, neurodegenerative disorders,
CC cardiomyopathies, viral infections and metabolic disorders
XX
SQ Sequence 2301 BP; 545 A; 568 C; 785 G; 403 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.427	Length:	2301
Score:	96.00	Matches:	29
Percent Similarity:	41.41%	Conservative:	24
Best Local Similarity:	23.66%	Mismatches:	51
Query Match:	13.43%	Indels:	24
DE:	3	Gaps:	3

US-10-005-907-2 (1-135) x AAA09303 (1-2301)

QY	14	GluAsnGlnLysLeuProLysGlyAsnProAspGluGluArgLysArgGlnGluMet	33
Dd	17	: : : : : : : : : : : : : : : : : : : :	
		GAGGACGAGGAGCAGGACGAGGACGAGGACGAGGAGGACGAGGAGGACGAG	76
QY	34	ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerThr	53
Dd	77	: : : : : : : : : : : : : : : : : : : :	
		GAGGACGAGGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	136
QY	54	SerAsnGlnGluAsnGlySerGlySerGluGluValCys-----	68
Dd	137		
		GAGGACGAGGACGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGC	196
QY	69	-----TyThrValIleAsn	73
Dd	197	AGGAACCGGAGCGGAGCAGTAGCTGGTGGCCACCATGCTGGATCACCATTCGAG	256
QY	74	HisIleProHisGlnArgSerSerLeuSerSerAsn---AspAspGlyTyrgluAsnIle	92
Dd	257	: : : : : : : : : : : : : : : : : : : :	
		GCGGTGAAGCGCAAGATCCAGGTTCCTGCAGCAGCAGGCAGATGATGCAGGAGGCGAGCT	316
QY	93	AspSerLeuThrArgLysVal-----ArgGlnPheArgGluArgSerGluThrGlu	109
Dd	317	: : : : : : : : : : : : : : : : : : : :	
		GAGCGCTCCAGCGAGAAGTTGAGGGAGAAAGCGGCCCGGAAACAGGCTGAGGCTGAG	376
QY	110	TyrAlaLeuLeuArgThrSerVal	117
Dd	377		
		GTGGCTCTCTTGAACCGTAGGATC	400

RESULT 15

ABL27495	
ID	ABL27495 standard; DNA; 1824 BP.
XX	
XX	
AC	ABL27495;
XX	
XX	
DT	26-MAR-2002 (first entry)
XX	
XX	
DE	Drosophila melanogaster genomic polynucleotide SEQ ID NO 33958.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical; gene; ds.
XX	
OS	Drosophila melanogaster.

XX
DN

FN
XX
WO200171042-A2.

27-SEP-2001

XX

PF 23-MAR-2001; 2001WO-US009231.

XX
PB 33-MAR 3000

PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000: 2000US-00634150

XX
XX

PA (PEKE) PE CORP NY.

[illegible]

PI
XX

WPI: 2001-656860/75

[illegible]

PT PT PT XX PS XX CC CC CC CC CC CC CC CC CC CC XX SQ

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

US-10-005-907-2 (1-135) x ABL27495 (1-1824)

Qy	3	AsnTyrLeuLeuArgLysLeuSerCys---	LeuGlyGluAenGlnLysLysProLysLys	21
Db	940	AACcAGgAgTTGGGGATCTGTCTCTCTCTCGAcAGcACCcGCcAGAAGgACCGCAAG	999	
Qy	22	-----GlyAsnProAspGlUGluArgLysArgInGluMet	33	
Db	1000	CTGCGCGGGAAATGGCAAAGATTTGCGCGATACACGGCCACGCCTGATGCGCGAGGAGTG	1059	
Qy	34	ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal-	50	
Db	1060	GCCCGGTACCAGAACAACTACGTCACTGCTGATTAAGCAGCAGAGCTAATCACCGAT	1119	
Qy	50	-----	50	
Db	1120	AATCTAGAGCTCAAGAGCTCTCCCTCTACTTCGACGAGAGCGTGCCTCCCATGTGGCGGCA	1179	
Qy	51	-----SerSerThrSerAsnGlnGluAsnGlySer	62	
Db	1180	AATGCGCTGTGCCCGGTTGCGGAGCAGCCACTAGGAATCGCGTTCAGATGACGGCGAT	1239	
Qy	63	GlySerGlu-----GluValCysTyrThrVallleAsnHisIleProHis	77	
Db	1240	GGTTCAGCTCCAGCAGCAATGCAGTAGAGCACCATCACCGCTTTGAGGAACATGCAGAG	1299	
Qy	78	GlnArgSer-----SerLeuSerSerAsnAspGlyTyrGlu	90	
Db	1300	CAGCGGCACTACCCcAGGATCTAGCTATGCCCCACCGCTGAACcACcAGCGCTTCAG	1359	
Qy	91	AsnIleAspSerLeuThrArgLysValArgGlnPheArgGlnArgSerGluThrGluTyr	110	
Db	1360	TATGTGCGTCACTGGAGCGCGCATCCAGCAACTTGGAGGAGGAGCGCCACCGCCACG	1419	
Qy	111	AlaLeuLeuArg	114	
Db	1420	GCCCATCTACAA	1431	

Search completed: November 19, 2004, 01:51:20
Job time : 430 secs

Db 449 CGGGAGCAGCGCTTCTCCAAAGTCAAAATTAGACTGTCTCTCTGAGCCTGTACTTCACCA 50
 QY 126 uHis 127
 Db 509 GCAC 512
 RESULT 13
 AAV20445
 ID AAV20445 standard; DNA; 2301 BP.
 AC AAV20445;
 DT 17-JUN-1998 (first entry)
 XX Human c-trk oncogene.
 DE
 DE
 XX Human; oncogene; proto-oncogene; neoplastic disease; anticancer; cancer;
 KW antisense oligonucleotide; c-trk; ds.
 OS Homo sapiens.
 XX
 XX US9734039-A.
 PN
 PD 31-MAR-1998.
 XX
 XX 15-SEP-1994; 94US-00306691.
 PF
 XX 15-SEP-1994; 94US-00306691.
 PR
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 PA
 XX Calabretta B, Skorski T;
 PI
 XX WPI; 1998-229882/20.
 DR
 XX Anticancer composition comprising two anti-sense oligo:nucleotide(s) -
 PT targeting cytoplasmic and nuclear oncogene(s).
 TT
 XX Claim 1; Col 99-102; 92pp; English.
 PS
 XX The present sequence represents an oncogene from the present invention.
 CC The present invention describes a composition which comprises two
 CC antisense oligonucleotides. The first oligonucleotide is specific for a
 CC cytoplasmic oncogene or proto-oncogene selected from ras, rat, EGF-1, c-
 CC fms, c-cro, c-kit, c-met, c-trk, c-sar, c-abl, bcr-abl, c-fgr and c-yes.
 CC The second oligonucleotide is specific for a nuclear oncogene or proto-
 CC oncogene selected from myc, jun, c-ets, c-fos, c-myb, B-myb, c-rel, c-
 CC vav, c-ski, c-spi, cyclin D1, PMW/RAR alpha, AML1/MTG8, E2A/prl and ALL
 CC 1/AF-4. The composition is used for treating cancer. The combination of
 CC antisense oligonucleotides has synergistically enhanced ability to
 CC inhibit growth of cancer cells
 XX
 SQ Sequence 2301 BP; 545 A; 568 C; 785 G; 403 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.427 Length: 2301
 Score: 96.00 Matches: 29
 Percent Similarity: 41.41% Conservative: 24
 Best Local Similarity: 22.66% Mismatches: 51
 Query Match: 13.43% Indels: 24
 DB: 2 Gaps: 3
 US-10-005-907-2 (1-135) x AAV20445 (1-2301)
 QY 14 GluAnGlnLysLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMet 33
 Db 17 GAGGAGCAGGAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGGAGGAGGAGGAG 13
 QY 34 ThrThrPheGluArgLeuGlnAspGlnAspLysLysSerGlnGluValSerThr 53
 Db 77 GAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 13
 QY 54 SerAsnGlnGluAsnGluAsnGlySerGlySerGluGluValCys----- 58


```

XX PD 23-SEP-1999.
XX PF
XX PP 18-MAR-1999; 99WO-US005804.
XX PR 19-MAR-1998; 98US-0078563P.
XX PR 19-MAR-1998; 98US-0078566P.
XX PR 19-MAR-1998; 98US-0078573P.
XX PR 19-MAR-1998; 98US-0078574P.
XX PR 19-MAR-1998; 98US-0078576P.
XX PR 19-MAR-1998; 98US-0078577P.
XX PR 19-MAR-1998; 98US-0078578P.
XX PR 19-MAR-1998; 98US-0078579P.
XX PR 19-MAR-1998; 98US-0078581P.
XX PR 01-APR-1998; 98US-0080312P.
XX PR 01-APR-1998; 98US-0080313P.
XX PR 01-APR-1998; 98US-0080314P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;
XX PF Wei Y, Endress GA, Duan RD, Kyaw H, Ebner R, Lafleur DW, Olsen HS;
XX PP Shi Y, Moore PA;
XX PR WPI; 1999-562050/47.
XX PR P-PSDB; AAY41387.
XX PT New isolated human genes, useful for diagnosis and treatment of e.g.
XX PT cancers, neurological disorders, immune diseases, inflammation or blood
XX PT disorders.
XX PS Claim 1; Page 346; 484pp; English.
XX CC This sequence represents a nucleic acid molecule which encodes a secreted
XX CC human protein. The gene number, and the clone it is derived from, are
XX CC detailed in the descriptor line. The gene can be used to generate fusion
XX CC proteins by linking to the gene to a human immunoglobulin Fc portion
XX CC (e.g. AA224802) for increasing the stability of the fused protein as
XX CC compared to the human protein only. The invention relates to 95 novel
XX CC genes and their fragments (nucleic acid sequences: AA224811-Z24907; amino
XX CC acid sequences AAY41308-Y41404) which are useful for preventing, treating
XX CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
XX CC pathological conditions can be diagnosed by determining the amount of the
XX CC new polypeptides in a sample or by determining the presence of mutations
XX CC in the new polynucleotides. Specific uses are described for each of the
XX CC 95 polynucleotides, based on which tissues they are most highly expressed
XX CC in (see AA224811 for described uses)
XX SQ Sequence 1892 BP; 461 A; 565 C; 472 G; 394 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.19 Length: 1892
Score: 98.00 Matches: 36
Percent Similarity: 38.26% Conservative: 21
Best Local Similarity: 24.16% Mismatches: 62
Query Match: 13.71% Indels: 30
DB: 2 Gaps: 4

US-10-005-907-2 (1-135) x AA224890 (1-1892)

QY 5 LeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLysLysGlyAsnPro 24
Db 572 CTCTCAATCTCTCTTGTCTGTCCTGGAAGCCAGGAGCCAGCCAGCAGCAGCAGGAG 631
QY 25 AspGluGluArgLysArgGlnGluMetThrPheGluArgLysLeuGlnAspGlnAsp 44
Db 632 CAAGGAGTGAGCAGCAGGAGGAGCCGACACAGAAACACAGCAGGAGGAGGAGGAGAA 691
QY 45 LysLysSerGlnGluValSerSerThrSerAsnGlnLysAsnGluAsnGlySerGlySer 64
Db 692 CAGGAGAGCCAGCAGCAGCAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 751
QY 65 GluGlu-----ValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80

752 AAGGAGGAGCGGAGGCTGTCTCTCAGCTGCAGACAGACTCAGAGCCCAAGTTTCACTCT 811
81 ---SerLeuSerSerAsnAspAspGlyTyr----- 89
812 GAATCTCTATCTTCTTAACCCCTTCTCTTCTCCCGGGTACGAGAAGTAGATCTACT 871
90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu--- 104
872 CCTATGATAATGGAGAACATCCAGGAGCTCATTCGATCAGCCAGGAATAAGATGAATG 931
105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
932 AATGAAATATATGATGAGAACTCTACTGTGAGAACCAAAACCTGTCAGCCTCTGCAG 991
115 ThrSerValSerArgProCysSerCys 123
992 CTGCCCCACACAGAGCCTTCTGTGTGC 1018

RESULT 7
ADA39940
ID ADA39940 standard; cDNA; 1892 BP.
XX AC ADA39940;
XX AC ADA39940;
XX DT 20-NOV-2003 (first entry)
XX DE Human secreted protein encoding cDNA.
XX KW Human; secreted protein; cancer; hyperproliferative disorder;
XX KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
XX KW anaemia; allergic reaction; asthma; cardiovascular disorder;
XX KW wound healing; cycostatic; immunosuppressive; nootropic; neuroprotective;
XX KW antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;
XX KW vulnery; cardiant; gene therapy; ss.
XX OS Homo sapiens.
XX PN WO2002102993-A2.
XX PD 27-DEC-2002.
XX PF 19-MAR-2002; 2002WO-US008123.
XX PR 21-MAR-2001; 2001US-0277340P.
XX PR 19-JUL-2001; 2001US-0306171P.
XX PR 13-NOV-2001; 2001US-0331287P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX PF WPI; 2003-175238/17.
XX PT New human secreted proteins and nucleic acid molecules, useful for
XX PT preparing a diagnostic or pharmaceutical composition for diagnosing,
XX PT preventing or treating cancer or other hyperproliferative disorder,
XX PT asthma, allergies or AIDS.
XX PS Claim 9; SEQ ID NO 322; 3205pp; English.
XX CC The invention relates to novel genes ADA39629-ADA40565 and proteins
XX CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,
XX CC treating or ameliorating medical conditions e.g. by protein or gene
XX CC therapy. The polypeptides, nucleic acid molecules, antibodies or their
XX CC fragments, and agonists or antagonists that bind to the polypeptide are
XX CC useful for preparing a diagnostic or pharmaceutical composition for
XX CC diagnosing or treating cancer or other hyperproliferative disorder. The
XX CC polypeptides and nucleic acid molecules are also useful for detecting,
XX CC preventing, diagnosing, prognosticating, treating or ameliorating cancer
XX CC or other hyperproliferative disorders including neoplasms, autoimmune
XX CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
XX CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic

```

CC encode. The polynucleotides and polypeptides are useful in diagnostics, forensics, gene mapping, medical imaging, identification of mutations responsible for genetic disorders or other traits, assessing biodiversity and producing many other types of data and products dependent on DNA and amino acid sequences. They are also useful for preventing, treating or ameliorating medical conditions, such as cancer, neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell disorders (e.g. osteoporosis, osteoarthritis, bone degenerative disorders, periodontal disease, liver fibrosis, infections (e.g. viral, fungal or bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis). Sequences ABX04986-ABX0511 represent human polynucleotides of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office

XX SQ Sequence 775 BP; 226 A; 202 C; 160 G; 187 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4, 7e-08 Length: 775
Score: 149.00 Matches: 37
Percent Similarity: 60.13% Conservatives: 28
Best Local Similarity: 34.26% Mismatches: 39
Query Match: 20.84% Indels: 4
DB: 8 Gaps: 4

US-10-005-907-2 (1-135) x ABX05108 (1-775)

QY 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGluVal 50
DB 329 AAAAAAATCACTCATTTTGAAGAAGG---CAAGATTCACAAACGAAATGAAGAATG 385

QY 51 SerSerThrSer---AsnGlnGluAsnGluAsnGlySerGlySerGluGluValCysTyr 69
DB 386 TCATCTACTCCATCCAGCAGGACAATGTTGACAGACTCTCTCAGAGAGCTGCTAT 445

QY 70 ThrValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspAspGly 88
DB 446 ACCCTCATCATCATCGGGTCTCTGTACAGGCCATCAGGAGACTCTCTGTAAGAGTAC 505

QY 89 TyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThr 108
DB 506 TATGAAATGTTCCCTGCAAGCTGAGACCCAGAGAGTCTTGGGAGGAACTGAGACT 565

QY 109 GluTyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHis 127
DB 566 GAGTATTCATCTCTACATATGCTTCTACAGACCCCGAGGATGCCCGATCCCGAAGAT 625

QY 128 AspTyrGluValValPheProHis 135
DB 626 GAATATGAACCTTCTCATGCTCTAC 649

RESULT 6
AAZ24890
ID AAZ24890 standard; DNA; 1892 BP.
XX AC AAZ24890;
XX DT 02-DEC-1999 (first entry)
XX DE Human secreted protein gene 80 clone HLMW54.
XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy; diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia; developmental abnormality; blood; deficiency; allergy; renal; ds; immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma; inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS; cognitive disorder; schizophrenia; prostate; obesity; thyroid; digestion; osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion; endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX OS Homo sapiens.
XX PN WO9947540-A1.

QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
DB 1156 AATGTGAGATGCAAAAGCCCAACAGAGAACATCCAGATCTGGATCACCATATCGCT 1215

QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
DB 1216 GAAGGGTCTTCTGCTCCATGGAAGAAATACTCATTTTGAAGAAGG---CAAGAT 1272

QY 43 GlnAspLysSerGlnGluValSerThrSer---AsnGlnGluAsnGluAsnGly 61
DB 1273 TCCCAAAACGAAATGAAGAATGTCATCTACTCCATCCAGCAGACAATGTTGACAG 1332

QY 62 SerGlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 80
DB 1333 ACCCTACAGAGAGCTGTGCTATACCTTCATCAATCATCGGGTCTCTGTACAGGCCA 1392

QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
DB 1393 TCAGGGAATCTGTGAGAGTACTATGAGATGTTCCCTGCAAGCTGAGAGACCCAGA 1452

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
DB 1453 GAGTCTCTGGGAGGAAGTGAAGTCTGATGATTCATCTCTANATATGCTTCTTACAGACCC 1512

QY 121 CysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
DB 1513 AGGCATGCCGATCCCAAGAGATGAATGAAGTCTTTCATGCTCTAC 1560

RESULT 5
ABX05108
ID ABX05108 standard; cDNA; 775 BP.
XX AC ABX05108;
XX DT 17-JAN-2003 (first entry)
XX DE Human novel polynucleotide #123.
XX KW Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer; neurodegenerative disorder; lymphoid cell disorder; osteoporosis; Parkinson's disease; Alzheimer's disease; bone degenerative disorder; osteoarthritis; periodontal disease; liver fibrosis; viral infection; fungal infection; bacterial infection; autoimmune disease; diabetes; atopic dermatitis.
XX OS Homo sapiens.
XX PN WO200274961-A1.
XX PD 26-SEP-2002.
XX PF 14-MAR-2002; 2002WO-US005109.
XX PR 15-MAR-2001; 2001US-00810173.
XX PA (HYSB-) HYSEQ INC.
XX PI Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
XX PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
XX PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX DR WPI; 2003-040556/03.
XX DR P-PSDB; ABU00030.
XX KW New isolated polypeptides and polynucleotides, useful for preventing, treating or ameliorating medical conditions, such as cancer, neurodegenerative disorders, lymphoid cell disorders, bone degenerative disorders, and infections.
XX FS Claim 1; SEQ ID NO 123; 235pp; English.
XX PN The invention relates to human polynucleotides and the polypeptides they

CC response in a patient. The protein can also serve as a target that
CC modulate gene expression or activity and as an antigen to raise
CC polyclonal or monoclonal antibodies. (II) is useful for identifying
CC agents that modulate expression of the protein or agents, such as
CC agonists or antagonists. The agonists or antagonists are useful for
CC modulating biological activity and function of (II) and thus are useful
CC for alleviating disease conditions such as allergic hypersensitivity,
CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis
XX
SQ Sequence 3762 BP; 1220 A; 672 C; 680 G; 1190 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.5e-75 Length: 3762
Score: 715.00 Matches: 135
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-005-907-2 (1-135) x ABN81319 (1-3762)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
DB 25 ATGGGAAATATCTCTCTCGAAACTCAGTTGCTGGGAGAGAAATCAAAAGAACGCCAAG 84

QY 21 LysGlyAsnProAspGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 85 AAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACATCACTTTGAAAGAAACAT 144

QY 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
DB 145 CAAGATCAAGATAAGAAAGCCAGAAAGTTTCTATCCACTTCTTAATCAGGAAACGAGAT 204

QY 61 GlySerGlySerGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 205 GGCAGTGGTTCGAGAGGTGCTACACTGTCAATTAATCACATCCCCCATGAGATCC 264

QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
DB 265 TCCTGAGCTCCAATGATGCTGATGAGAACATTTGACTCCCTCAAGAAAGTGAGA 324

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
DB 325 CAGTTTAGAGAAAGGTCAGACAGACAGATATGCCCTCTTAGGACTTCTGTAGAGCCT 384

QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 385 TGTTCTCGCACCATGAGCATGATTATGAAGTTGTGTTCACAC 429

RESULT 3
ACC59951
ID ACC59951 standard; cDNA; 732 BP.
XX
AC ACC59951;
XX
DT 07-JUL-2003 (first entry)
XX
DE Human IRAP-28 encoding cDNA SEQ ID NO:63.
XX
KW Human; immune response associated protein; IRAP; anti-HIV; anti-allergic;
KW anti-inflammatory; antianemic; antiparkinsonian; neurotropic; antithyroid;
KW anticonvulsant; antiarteriosclerotic; antiasthmatic; immunosuppressive;
KW cytotatic; hepatotropic; dermatologic; antidiabetic; nephrotropic;
KW antitumor; thymimetic; neuroprotective; osteopathic; antiarthritic;
KW antiparasitic; antihelminthic; antipsoptic; uropathic; protozoacide;
KW antirheumatic; haemostatic; antibacterial; virucide; ophthalmological;
KW fungicide; gene therapy; immune system disorder; neurological disorder;
KW developmental disorder; muscle disorder; cell proliferative disorder;
KW infection; gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH

CDS 134..670
/*tag= a
/product= "IRAP-28"

WO2003025542-A2.
27-MAR-2003.
19-SEP-2002; 2002WO-US029979.
21-SEP-2001; 2001US-0324034P.
05-OCT-2001; 2001US-0327395P.
12-OCT-2001; 2001US-0328923P.
19-OCT-2001; 2001US-0342810P.
09-NOV-2001; 2001US-0344468P.
21-NOV-2001; 2001US-0332140P.
07-DEC-2001; 2001US-0340282P.
09-JAN-2002; 2002US-0347693P.
20-FEB-2002; 2002US-0358279P.
01-MAR-2002; 2002US-0361088P.
15-MAR-2002; 2002US-0364494P.
10-MAY-2002; 2002US-0379876P.
11-JUN-2002; 2002US-0389180P.
(INCY-) INCYTE GENOMICS INC.
Ho A, Baughn MR, Becha SD, Burford N, Elliott VS, Emerling BM;
Forsythe IU, Gorvad AE, Griffin JA, Hafalia AJA, Honchell CD;
Ison CH, Burdill JD, Blake JJ, Lal PG, Lee EA, Marquis JP;
Lehr-Mason PM, Lee S, Sprague WW, Swarnakar A, Tang YT, Tran B;
Tran UK, Bhatia U, Chawla NK, Warren BA, Zheng W, Xu Y, Yue H;
WPI: 2003-363161/34.
P-PSDB; ABR43232.
New human immune response associated proteins and polynucleotides, useful
for diagnosing, treating or preventing immune system disorders, e.g. AIDS
or anemia, cell proliferative disorders, e.g. cancer, or neurological
disorders.
Claim 5; Page 208; 213pp; English.
ACC59924 to ACC59958 encode the human immune response associated proteins
given in ABR43205 to ABR43239, designated IRAP-1 to IRAP-29 (I). (I) have
anti-HIV, anti-allergic, anti-inflammatory, antianemic, antiparkinsonian,
neurotropic, anticonvulsant, antiarteriosclerotic, antiasthmatic, antitumor,
immunosuppressive, antithyroid, cytotatic, hepatotropic, dermatologic,
antidiabetic, nephrotropic, thymimetic, neuroprotective, osteopathic,
antiarthritic, antiparasitic, antihelminthic, antipsoptic, uropathic,
ophthalmological, antirheumatic, haemostatic, antibacterial, virucide,
protozoacide and fungicide activities, and can be used in gene therapy.
Human IRAP polynucleotides, agonists and antagonists are useful for
diagnosing, treating or preventing disorders associated with aberrant
expression of IRAP, such as immune system disorders (e.g. AIDS, asthma,
allergies, autoimmune thyroiditis, contact dermatitis, Crohn's disease,
diabetes mellitus, glomerulonephritis, Good pasture's syndrome, gout,
Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis),
neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or
epilepsy), developmental disorders (e.g. renal tubular acidosis, anaemia
or mental retardation), muscle disorder (e.g. cardiomyopathy, myocardiitis
or dermatolysisitis), cell proliferative disorders (e.g. arteriosclerosis,
atherosclerosis, cirrhosis, hepatitis, psoriasis, paroxysmal nocturnal
haemoglobinuria, polycythaemia vera, psoriasis, primary
thrombocytopaenia or cancer), or viral, bacterial, fungal, parasitic,
protozoan or helminthic infections
SQ Sequence 732 BP; 213 A; 178 C; 178 G; 163 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.55e-10 Length: 732
Score: 169.50 Matches: 47

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Ma Y, Asundi V, Wang Z, Meng G;
PI Haley-Vicente D, Dmanac RT;
XX WPI; 2003-371981/35.
DR P-PSDB; ABC31731.
XX
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
XX Claim 1; SEQ ID NO 842; 1185pp; English.
XX
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 93% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 667 BP; 216 A; 147 C; 154 G; 150 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9 55e-76 Length: 667
Score: 715.00 Matches: 135
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-005-907-2 (1-135) x ADC30760 (1-667)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAsnGlnLysLysProLys 20
Db 186 ATGGGAAATTTATCTCTCGAAACTCAGTTGCTGGGAGAGATCAAGAAGACCCAG 245
QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
Db 246 AAGGAAACCCAGATGAGGAAAGAAACCGGAGAAATCACTACATTTGAAGAAACTT 305
QY 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
Db 306 CAGATCATAGATAGAAACCCAGAGAGTTTCTACCTCTCTATCAGGAAACGAGAAT 365
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80

Db 366 GGCAGTGTGTTCTGAAGAAGTGTGTACACTGTCTATTAATCACATCCCCCATCAGATCC 425
QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
Db 426 TCCCTGAGCTCCCAATGATGATGGCTATGAGAACATTGACTCCCTCCACAGGAAAGTGA 485
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
Db 486 CAGTTTGAAGAAAGTTCAGAGACAGAAATATGCCCTCTTAGGACTCTCTGTAGTAGGCCT 545
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
Db 546 TGTCTCTGCCCATGAGCATGATTATGAAGTTGTGTTCACAC 590

RESULT 2
ABN81319
ID ABN81319 standard; cDNA; 3762 BP.
XX
AC ABN81319;
XX
DT 30-AUG-2002 (first entry)
DE Human mast cell related gene MC1 SEQ ID NO 1.
DE
KW Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
KW vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma;
KW gene; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 25..432
FT /*tag= a
FT /product= "MC1"
XX
PN W0200246389-A2.
PD
PD 13-JUN-2002.
XX
XX 07-DEC-2001; 2001WO-US046180.
PF
XX 08-DEC-2000; 2000US-0251835P.
PR
PR 14-MAR-2001; 2001US-0275479P.
PR 28-MAR-2001; 2001US-0279115P.
PR 02-APR-2001; 2001US-0280143P.
XX
XX (UNIO) UCB SA.
XX
XX Nocka K, Pirozzi G, Einstein R;
PI
XX WPI; 2002-508560/54.
DR P-PSDB; ABN77569.
XX
PT Novel isolated nucleic acids that are differentially expressed in mast
PT cells in patients with allergic hypersensitivity, encoding proteins
PT associated with mast cell regranulation and allergic hypersensitivity.
XX
XX Claim 1; Page 95-97; 119pp; English.
XX
XX The invention relates to isolated nucleic acid (ABN81319-ABN81324),
CC corresponding to genes differentially expressed in mast cells following
CC activation or in patients with allergic hypersensitivity disease, (I)
CC that encodes proteins (ABN77569-ABN77575) (II) or a protein fragment of
CC (II) if at least 6 amino acids. (II) is useful for identifying binding
CC partners. (I) or (II) is useful for diagnosing or treating a disease
CC state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
CC urticaria or atopic dermatitis or mastocytosis) in a subject which
CC involves determining the level of expression of (I) or (II). A computer
CC system, comprising a database containing information identifying the
CC expression level in a tissue or at least one mast cell of (I), is useful
CC for presenting information to identify the relative expression level of
CC (I). (II) is used as a marker to detect, diagnose or identify an allergic

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 18, 2004, 21:26:34 ; Search time 424 Seconds

(without alignments)
1671.396 Million cell updates/sec

Title: US-10-005-907-2

Perfect score: 715
Sequence: 1 MGNVLLRLKSLGENQKPK.....SVSRPCSTHEHDYEVVFP 135

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO spool/US10005907/runat_16112004_140457.6157/app_query.fasta_1.327
-DB=N Geneseq 23Sep04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10005907@cgn_1.1.470 @runat_16112004_140457.6157 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	715	100.0	667	10	ADC30760 Human nov
2	715	100.0	3762	6	ABN81319 Human mas
3	169.5	23.7	732	8	ACC59951 Human IRA
4	159	22.2	1643	6	AD199791 Human sec
5	149	20.8	775	8	ABX05108 Human nov
6	98	13.7	1892	2	AZZ24890 Human sec

7	98	13.7	1892	8	ADA39940	Human sec
8	98	13.7	1892	10	ADC73578	Human sec
9	98	13.7	1892	10	ADD37623	Human sec
10	98	13.7	1892	10	ADA56130	Gene enco
11	97	13.6	3957	12	ADQ24942	Human sof
12	96.5	13.5	584	10	ADQ24942	Rat gene
13	96	13.4	2301	2	AAV20445	Human c-t
14	96	13.4	2301	3	AAA09303	Human Trk
15	93	13.0	1824	4	ABL27495	Drosophil
16	92	12.9	654	5	AA571153	DNA enco
17	92	12.9	654	5	AA575467	DNA enco
18	92	12.9	654	5	AA569553	DNA enco
19	91.5	12.8	2862	12	ADO00884	Human hom
20	91.5	12.8	3182	5	ABX711302	Human kid
21	91.5	12.8	3186	12	ADQ24514	Human sof
22	91.5	12.8	5109	10	ADF81540	Leukaemia
23	91.5	12.8	5261	5	ADL45687	Human ova
24	91.5	12.8	7488	6	ABL33112	Human imm
25	90.5	12.7	651	12	ADJ67693	Kifc DNA
26	90	12.6	1671	4	AAH99858	Human pro
27	90	12.6	1899	2	AA506785	Human adu
28	90	12.6	1912	3	AA250927	Human pro
29	90	12.6	7992	6	ABK32895	DNA enco
30	90	12.6	8083	5	AA570692	DNA enco
31	90	12.6	8257	9	ACA90154	Human cdn
32	90	12.6	8257	12	ADQ17931	Human sof
33	90	12.6	8491	10	ADI02719	Human cdn
34	90	12.6	8493	8	ACA03930	cDNA down
35	90	12.6	8503	4	AA158253	Human pol
36	90	12.6	8503	5	ADQ98460	DNA enco
37	90	12.6	8503	9	ADB48220	Novel hum
38	90	12.6	8517	12	ADQ22543	Human sof
39	90	12.6	8527	2	AA160039	Human pol
40	89.5	12.5	1027	4	AA84690	CRCALD05
41	89.5	12.5	1027	6	ABK47580	cDNA enco
42	89.5	12.5	1071	4	ABA83301	Human sec
43	89.5	12.5	1071	9	ACH04802	Novel hum
44	89.5	12.5	1071	9	ACD44612	Human cdn
45	89.5	12.5	3719	10	ADI40373	Human pur

ALIGNMENTS

RESULT 1
ADC30760
ID ADC30760 standard; cDNA; 667 BP.
XX
AC ADC30760;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel cDNA sequence, SEQ ID NO:842.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 1; gene; ss.
XX
OS Homo sapiens.
XX
FN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.

QY 36 PheGluArgLysLeuGlnAspGlnAspLysSerGlnGluValSerSerThrSerAsn 55
 Db 64 CGTGTGAAGAAGTACAAAGAGCAAGCAAAACAG---CAGAAAGTGGAGTTTCGGAAAAGG 120
 QY 56 GlnGluAsnGlu-----AsnGlySerGlySerGluGluValCysTyrThr 70
 Db 121 ATGGAGAAAGAGGTGCTGATTTCATCCAGGACAGTGGACAGGTCAAGAAAAAGTTTCAG 180
 QY 71 ValIleAsnHisIleProHisGlnArgSer----- 80
 Db 181 CCTATGAACAAGATA-----GAGCGGAGCATACTACATGATGTGGTAGAGTGGCTGGC 234
 QY 81 -----SerLeuSerSerAsnAspAsp----- 87
 Db 235 CTCACATCCCTCTCTCTTTGGAGAGATGATGACTGTGCTGCTATGTCATGATCTTCAAAAAG 294
 QY 88 -----GlyTyrGluAsnIleAspSerLeu----- 95
 Db 295 GAGTTTGACCTTCAGATGAAGAGCTAGACTCTCTACCGTCATGGAGAGGAGTGGACCCC 354
 QY 96 -----ThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyr 110
 Db 355 CAGAAAGCTGAGGAGAGCGGAAGCTAAAGGAGCTGGCTCAGAAAGCAGGAGGAGGCA 414
 QY 111 AlaLeuLeuArgThrSerValSerArgProCysSerCysThrHisGluHisAspTyrGlu 130
 Db 415 GCGCAGCAGGACCTGGGTTGTGAGTCTCTCCAGC-----GACTACAAG 459
 QY 131 ValValPheProHis 135
 Db 460 GACAAGTATAGCCAT 474

RESULT 15

US-10-437-963-64985
 ; Sequence 64985, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Bing
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 64985
 ; LENGTH: 2693
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_66077C.1
 US-10-437-963-64985

Alignment Scores:
 Pred. No.: 0.373 Length: 2693
 Score: 90.50 Matches: 35
 Percent Similarity: 39.86% Conservative: 20
 Best Local Similarity: 25.36% Mismatches: 62
 Query Match: 12.66% Indels: 21
 DB: 17 Gaps: 4

US-10-005-907-2 (1-135) x US-10-437-963-64985 (1-2693)

QY 12 LeuGlyGluAsnGlnLysProLysLysGlyAsnProAspGluGluArgLysArgGln 31
 Db 705 TTGGAAAGACAAATGATGATCAAGAAAGAAAGATTCAAGAACCCGAGGTTGAAAGAGCA 764

QY 32 GluMetThrThrPheGlu-----ArgLysLeuGlnAspGlnAsp----- 44
 Db 765 GAAGAAACAAACCAAGGATGGAAGTTGTTTGATCGCTTCTCAACAAGAAAGAAATTTGGTG 824
 QY 45 -----LysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
 Db 825 GCATTTCTTATAGAAAAATCGTGGATGAAAAAATCACTTTTGAACAAGAAAAAGAAAG 884
 QY 61 GlySerGlySerGluGluVal-----CysTyrThrValIleAsn 73
 Db 885 GTAGAAGGCTCTGAAGAAACAAAGAACCAAGTAAAGACAGCTTATGATGTCGAGATG 944
 QY 74 HisIleProHisGlnArgSerSerLeuSerSerAsnAspGlyTyrGluAsnIleAsp 93
 Db 945 GACGTC-----CAAAAATCATTTGAAACTGATGAGAGCTCAAAAGGACACTTGT 995
 QY 94 SerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeu 113
 Db 996 GATTTGGAAGCACCGATATACCAAGATAATGCCCAGATGAAATTTCTAGCAAACTGATA 1055
 QY 114 ArgThrSerValSerArgProCysSerCysThrHisGluHisAspTyrGluVal 131
 Db 1056 TCAGCCAAAGGACGCTGATCCGATCAAAAAAATGGAAGCTAGAGACTTTGATTTG 1109

Search completed: November 19, 2004, 05:28:12
 Job time : 1530 secs

LOCATION: 1, 2, 3, 4, 5, 6, 5250, 5251, 5252, 5253, 5254, 5255, 5256,
LOCATION: 5257, 5258, 5259, 5260, 5261
OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19577

Alignment Scores:
Pred. No.: 0.702 Length: 5261
Score: 91.50 Matches: 32
Percent Similarity: 47.90% Conservative: 25
Best Local Similarity: 26.89% Mismatches: 45
Query Match: 12.80% Indels: 17
DB: 10 Gaps: 4

US-10-005-907-2 (1-135) x US-09-814-353-19577 (1-5261)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
DB 610 ATCGGCCAGTGTGTTGAAGAAGAACAGACCCCTAACCGGAGGAAACGAGCTGTCTGGAG 669

QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 670 GAG-----CAGGTGAACACATCAGGAGGAGGTGTCTCAGCTCCGGCATGAGCTG 720

QY 41 GlnAspGlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
DB 721 TCCATGAGAGATGAGCTGCTTCACTTACACAGCGCTCGGAGGAG-----768

QY 61 GlySerGlySerGluGluValCysTyrThrValLeuAsnHisTleProHisGlnArgSer 80
DB 769 ---AGTGAGCCGAGTCCGTTTCTCAACC-----CCGTTGAGAGGAAT 810

QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnLeuAspSerLeuThrArgLysValArg 100
DB 811 GAGTCGCTCTCTCCAGTCCAGAAATTTTCACTTGGATTCTCTCAAGAGAGCTGAA 870

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
DB 871 GACCTTGAAGAG-----GAGATGTTGTACTTCGATCCGAGGCCACCCAG 915

RESULT 13
US-10-311-455-1085/c
Sequence 1085, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determination of Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
PRIOR FILING DATE: 2002-12-16
PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1085
LENGTH: 7488
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1085
Alignment Scores:
Pred. No.: 1.15 Length: 7488
Score: 91.50 Matches: 29
Percent Similarity: 44.22% Conservative: 36
Best Local Similarity: 19.73% Mismatches: 51
Query Match: 12.80% Indels: 31

DB: 15 Gaps: 5
US-10-005-907-2 (1-135) x US-10-311-455-1085 (1-7488)

QY 14 GluAsnGlnLysLys-----Pro 19
DB 6080 GAAAAATACATAAAAAATACGACTCACTTCTTCGCGCTCGACAAAATACACTACTATAA 6021

QY 20 LysLysGlyAsnProAspGluGluArgLysArgGlnGlu-----MetThrThr 35
DB 6020 AAACAAATACCCCGATAAAAAATAAATAAACACACCCCTAACTCTCTACTACTAAC 5961

QY 36 PheGluArg-----LysLeuGlnAspGlnAspLysLysSer 47
DB 5960 TTTAAAAATCTAACCCAAAATCTAAAAACCCCTTAAAAATACGAAAAAACTTAAAAATCT 5901

QY 48 GlnGluValSerSerThrSerAsnGlnGluAsnGluAsnGlySerGlySerGluGluVal 67
DB 5900 AAAAACCCCTTACCAACGAAATACCCCTAAAAATAACCAAAAAATAACCTCTCGCCCTT 5841

QY 68 CysTyrThr---ValLeuAsnHisTleProHisGlnArgSerSerLeuSerSerAsnAsp 86
DB 5840 CTAAATTCCTTCTACTCACTATCTCGATACCCACGTAATAAATAAATAAATAAATAAC 5781

QY 87 AspGlyTyrGluAsnLeuAspSerLeuThrArgLysValArgGlnPheArgGluArgSer 106
DB 5780 CACAATATCTATAAAATTCATAATATTTTAAAAAATCCGTTCTTACTTAAACACAAACGAA 5721

QY 107 GluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCysSerCysThrHisGlu 126
DB 5720 AAAACGACCCACCACTCGAAAA-----AAACCAATAATAACACACGTTTC 5673

QY 127 HisAspTyrGluValValPhe 133
DB 5672 CAGATAAAAAATAAATCTAT 5652

RESULT 14
US-10-261-175A-9
Sequence 9, Application US/10261175A
Publication No. US20040038222A1
GENERAL INFORMATION:
APPLICANT: DIETRICH, WILLIAM
APPLICANT: WATERS, JAMES W.
TITLE OF INVENTION: ANTHRAX SUSCEPTIBILITY GENE
FILE REFERENCE: 56491(71250)
CURRENT APPLICATION NUMBER: US/10/261,175A
CURRENT FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: 60/325,864
PRIOR FILING DATE: 2001-09-29
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 651
TYPE: DNA
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Kiflc nucleic acid sequence

US-10-261-175A-9
Alignment Scores:
Pred. No.: 0.0514 Length: 651
Score: 90.50 Matches: 39
Percent Similarity: 36.97% Conservative: 22
Best Local Similarity: 23.64% Mismatches: 51
Query Match: 12.66% Indels: 53
DB: 16 Gaps: 7

US-10-005-907-2 (1-135) x US-10-261-175A-9 (1-651)
QY 16 GlnLysLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThr 35
DB 4 GAGAAACACCCAGCCTCGGTGACCGAGAGTCCGCGCAAGGCGGAGACGAGTCC 63


```

; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1552
; LENGTH: 5109
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_014965
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1552

```

```

Alignment Scores:
Pred. No.: 0.673 Length: 5109
Score: 91.50 Matches: 32
Percent Similarity: 47.90% Conservative: 25
Best Local Similarity: 26.89% Mismatches: 45
Query Match: 12.80% Indels: 17
DB: 15 Gaps: 4

```

```

US-10-005-907-2 (1-135) x US-10-172-118-1552 (1-5109)

Qy 1 MetGlyAsnTyrLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 610 ATCGGCCAGTCGTTGTTGAAGAAGAACAGACCCTAACCCGAGAGGAACGAGCTGCTGGAG 669

Qy 21 LysGlyAsnProAspGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 670 GAG-----CAGGTGGAACACATCAGGAGGAGGTGCTCAGCTCCGCGCATGAGCTG 720

Qy 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 721 TCCATGAAGGATGAGCTGCTTACAGTCTTACACAGCGCTCGCGAGGAG----- 768

Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 769 ---AGTGAGCCGAGTCGTTGCTCAACC-----CCGTTGAAGAGGAAT 810

Qy 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 811 GAGTCGTCCTCCTCAGTCAGCAATTAATTTCAATTGGATTCTCTTCAAAAGAGCTGAAA 870

Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 871 GACCTTGAAGAG-----GAGAAATGTTGTACTTCGATCCGAGGCCAGCCAG 915

```

```

RESULT 11
US-10-342-887-1552
; Sequence 1552, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1552

```

```

; LENGTH: 5109
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1552

```

```

Alignment Scores:
Pred. No.: 0.673 Length: 5109
Score: 91.50 Matches: 32
Percent Similarity: 47.90% Conservative: 25
Best Local Similarity: 26.89% Mismatches: 45
Query Match: 12.80% Indels: 17
DB: 16 Gaps: 4

```

```

US-10-005-907-2 (1-135) x US-10-342-887-1552 (1-5109)

Qy 1 MetGlyAsnTyrLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 610 ATCGGCCAGTCGTTGTTGAAGAAGAACAGACCCTAACCCGAGAGGAACGAGCTGCTGGAG 669

Qy 21 LysGlyAsnProAspGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 670 GAG-----CAGGTGGAACACATCAGGAGGAGGTGCTCAGCTCCGCGCATGAGCTG 720

Qy 41 GlnAspGlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 721 TCCATGAAGGATGAGCTGCTTACAGTCTTACACAGCGCTCGCGAGGAG----- 768

Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 769 ---AGTGAGCCGAGTCGTTGCTCAACC-----CCGTTGAAGAGGAAT 810

Qy 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 811 GAGTCGTCCTCCTCAGTCAGCAATTAATTTCAATTGGATTCTCTTCAAAAGAGCTGAAA 870

Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 871 GACCTTGAAGAG-----GAGAAATGTTGTACTTCGATCCGAGGCCAGCCAG 915

```

```

RESULT 12
US-09-814-353-19577
; Sequence 19577, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19577
; LENGTH: 5261
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

Db 991 TCATCCGGTCAACACAAATGCTTATGCAAGAGGTAAACAAAGCATTCCTCT 1050
 Qy 79 ArgSerSerLeuSerSerAsp-----AspGlyTyrGluAsnIleAspSerLeuThr 96
 Db 1051 AGTGATGAAGTAGTAGTATGATGATGAGGTGAGGAAAGCCCTCTCTAGATGAGCTTCG 1110
 Qy 97 ArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSer 116
 Db 1111 GAAGCGGTAGAATTTTCCAGATGTTTGCCTAAGCAAAAGCTCAACTTAACCTTG 1170
 Qy 117 ValSerArgProCysSerCysThrHisGluHisAspTyrGlu 130
 Db 1171 AAAATAAGTTGGTTAGC-----TCTCAAAATGATTACAA 1206
 RESULT 8
 US-10-115-831-54
 ; Sequence 54, Application US/10115831
 ; Publication No. US20030219743A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and
 ; TITLE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 792CIP2ADIV
 ; CURRENT APPLICATION NUMBER: US/10/115,831
 ; CURRENT FILING DATE: 2002-04-02
 ; PRIOR APPLICATION NUMBER: 09/667,298
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: 09/577,408
 ; PRIOR FILING DATE: 2000-05-18
 ; NUMBER OF SEQ ID NOS: 178
 ; SOFTWARE: pt_FL_genes Version 2.0
 ; SEQ ID NO 54
 ; LENGTH: 6324
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(3573)
 US-10-115-831-54
 Alignment Scores:
 Pred. No.: 0.78 Length: 6324
 Score: 92.00 Matches: 27
 Percent Similarity: 38.60% Conservative: 17
 Best Local Similarity: 23.68% Mismatches: 50
 Query Match: 12.87% Indels: 20
 DB: 15 Gaps: 3
 US-10-005-907-2 (1-135) x US-10-115-831-54 (1-6324)
 Qy 14 GluAsnGlnLysProLysGlyAsnProAspGluGluArgLysArgGlnGluMet 33
 Db 337 GACAAATGACGAGCCAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 396
 Qy 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGluValSerSerThr 53
 Db 397 GAGGAGATCGAGGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAGGAGGATGAA 456
 Qy 54 SerAsnGlnGluAsnGluAsnGlySerGlySerGluGluVal----- 67
 Db 457 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 516
 Qy 68 -----CysTyrThrValIleAsnHisIleProHisGlnArgSerSerLeuSerSerAsn 85
 Db 517 ATGAAATGTCACAAAT-----ACTCGAATAATGCAACACACACAGAAAGGATGATAACAAT 570
 Qy 86 AspAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal----- 99
 Db 571 AATGACCAATATGACAAATACGATGAATGTTGCCCAAGTCAATGTTAAACCTCGGCAAA 630

Qy 100 -----ArgGlnPheArgGluArgSerGluThrGlu 109
 Db 631 ATCGCTGAGGATGACAGCTACCGGCCAGGAGTGAAGTCAAGAA 672
 RESULT 9
 US-10-263-929-22
 ; Sequence 22, Application US/10263929
 ; Publication No. US20040067535A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kim, Jaeseob
 ; APPLICANT: Galant, Ron
 ; TITLE OF INVENTION: Alzheimer's Disease Linked Genes
 ; FILE REFERENCE: LSD-07417
 ; CURRENT APPLICATION NUMBER: US/10/263,929
 ; CURRENT FILING DATE: 2002-10-03
 ; NUMBER OF SEQ ID NOS: 213
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 22
 ; LENGTH: 2862
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-263-929-22
 Alignment Scores:
 Pred. No.: 0.3 Length: 2862
 Score: 91.50 Matches: 32
 Percent Similarity: 47.90% Conservative: 25
 Best Local Similarity: 26.89% Mismatches: 45
 Query Match: 12.80% Indels: 17
 DB: 16 Gaps: 4
 US-10-005-907-2 (1-135) x US-10-263-929-22 (1-2862)
 Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
 Db 394 ATCGGCCAGTCTGTTGAAGAAGAACAAAGACCCCTAACCCAGAGGAGGAGCTCTCAGCTCCGCGATGAGCTG 453
 Qy 21 LysGlyAsnProAspGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
 Db 454 GAG-----CAGTGGAAACACATCAGGAGGAGGAGTCTCAGCTCCGCGATGAGCTG 504
 Qy 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
 Db 505 TCCATGAGGATGAGCTGCTTCACTTCTACACAGCGCTCGGAGGAG----- 552
 Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 Db 553 ---AGTGAGCCCGAGTCCGTTTGTCTCAACC-----CCGTTGAAGAGGAGAT 594
 Qy 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 Db 595 GAGTCGCTCTCTCAGTCCAGATTAATCTTTCATTTGATTCCTTCAAGAGAGCTGAA 654
 Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
 Db 655 GACCTTGAAGAG-----GAGAAATGTTGACTTCGATCCGAGGAGCCAGCCAG 699
 RESULT 10
 US-10-172-118-1552
 ; Sequence 1552, Application US/10172118
 ; Publication No. US20030224374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Chris
 ; APPLICANT: Van 't Veer, Laura
 ; APPLICANT: Van de Vijver, Marc
 ; APPLICANT: Bernards, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-175-999

US-09-397-945-90

Alignment Scores:

Pred. No.: 0.0237 Length: 1892
Score: 98.00 Matches: 36
Percent Similarity: 38.26% Conservatives: 21
Best Local Similarity: 24.16% Mismatches: 62
Query Match: 13.71% Indels: 30
DB: 10 Gaps: 4

US-10-005-907-2 (1-135) x US-09-397-945-90 (1-1892)

Qy 5 LeuLeuArgLysLeuSerCysLeuGluAsnGlnLysProLysLysGlyAsnPro 24
Db 572 CTCCTACAAATCCTCTGCTGCGAAGCCAGGAGCAAGCCAGAGCAGCAGAGGAG 631
Qy 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
Db 632 CAAGGAGTGGACACAGGAGGAGCGGACACAAAGAACACACAGCAGGAAGAGGGGCGAGAA 691
Qy 45 LysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnGlySerGlySer 64
Db 692 CAGGAAGAGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 751
Qy 65 GluGlu-----ValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db 752 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811
Qy 81 ---SerLeuSerSerAsnAspAspGlyTyr----- 89
Db 812 GAATCTCTATCTTCTTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 871
Qy 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu----- 104
Db 872 CCTATGATAATGAGAGAACATCCAGGAGCTCATTCGATCAGCCAGGAGAAATAGATGAATG 931
Qy 105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
Db 932 AATGAATAATATGATGAGAACTCTACTGGAGAAACCAAAACCTCGCAGCTCTCTGCGAG 991
Qy 115 ThrSerValSerArgProCysSerCys 123
Db 992 CTGCCCCACACAGAGCTTGCTGTGTC 1018

RESULT 4

US-10-653-595-90
; Sequence 90, Application US/10653595
; Publication No. US20040048304A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 95 Human secreted proteins
; FILE REFERENCE: P2027P1C1
; CURRENT APPLICATION NUMBER: US 09/397945
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: PCT/US99/05804
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/078,566
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,576
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,573
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,574
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,579
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/080,314
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080,312
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/078,578

; PRIOR FILING DATE: 1998-03-19
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 470
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 1892
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-653-595-90

Alignment Scores:

Pred. No.: 0.0237 Length: 1892
Score: 98.00 Matches: 36
Percent Similarity: 38.26% Conservatives: 21
Best Local Similarity: 24.16% Mismatches: 62
Query Match: 13.71% Indels: 30
DB: 16 Gaps: 4

US-10-005-907-2 (1-135) x US-10-653-595-90 (1-1892)

Qy 5 LeuLeuArgLysLeuSerCysLeuGluAsnGlnLysProLysLysGlyAsnPro 24
Db 572 CTCCTACAAATCCTCTGCTGCGAAGCCAGGAGCAAGCCAGAGCAGCAGAGGAG 631
Qy 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
Db 632 CAAGGAGTGGACACAGGAGGAGCGGACACAAAGAACACACAGCAGGAAGAGGGGCGAGAA 691
Qy 45 LysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnGlySerGlySer 64
Db 692 CAGGAAGAGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 751
Qy 65 GluGlu-----ValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db 752 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811
Qy 81 ---SerLeuSerSerAsnAspAspGlyTyr----- 89
Db 812 GAATCTCTATCTTCTTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 871
Qy 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu----- 104
Db 872 CCTATGATAATGAGAGAACATCCAGGAGCTCATTCGATCAGCCAGGAGAAATAGATGAATG 931
Qy 105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
Db 932 AATGAATAATATGATGAGAACTCTACTGGAGAAACCAAAACCTCGCAGCTCTCTGCGAG 991
Qy 115 ThrSerValSerArgProCysSerCys 123
Db 992 CTGCCCCACACAGAGCTTGCTGTGTC 1018

RESULT 5

US-10-242-943-3
; Sequence 3, Application US/10242943
; Publication No. US20030087412A1
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Schulz, Vincent P.
; APPLICANT: Yang, MeiJa
; TITLE OF INVENTION: Nlk1 PROTEIN AND Nlk1 PROTEIN COMPLEXES
; FILE REFERENCE: 15966-521 Nlk1 protein complexes
; CURRENT APPLICATION NUMBER: US/10/242,943
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US/09/167,206
; PRIOR FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: November 19, 2004, 02:51:09 ; Search time 1525 Seconds
(without alignments)
478.354 Million cell updates/sec

Title: US-10-005-907-2
Perfect score: 715
Sequence: 1 MGNVLLRLKSLGKQKPK.....SVSPSCSTHEDYEVVFP 135

Scoring table:
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3627888 seqs, 2701811610 residues
Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_model/US10005907/runat_16112004_140459_6250/app_query.fasta_1.327
-DB=published_applications NA -QFWT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITIS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR_SCORE=pct -THR_MAX=100
-NAXLEN=2000000000 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	3762	15	US-10-005-907-1
2	98	13.7	1173	18	US-10-425-115-147529
3	98	13.7	1892	10	US-09-397-945-90
4	98	13.7	1892	16	US-10-653-595-90
5	96	13.4	2301	14	US-10-242-943-3
6	93	13.0	1903	16	US-10-424-599-8569
7	92	12.9	1941	18	US-10-425-115-147532
8	92	12.9	6324	15	US-10-115-831-54
9	91.5	12.8	2862	16	US-10-263-929-22
10	91.5	12.8	5109	15	US-10-172-118-1552
11	91.5	12.8	5109	16	US-10-342-887-1552
12	91.5	12.8	5261	10	US-09-814-353-19577
13	91.5	12.8	7488	15	US-10-311-455-1085
14	90.5	12.7	651	16	US-10-261-175A-9
15	90.5	12.7	2693	17	US-10-437-963-64985
16	90.5	12.7	1173	13	US-10-027-632-10372
17	90	12.6	1173	15	US-10-027-632-10372
18	90	12.6	1671	16	US-10-296-115-693
19	90	12.6	1899	10	US-09-746-783-77
20	90	12.6	7992	10	US-09-893-519A-140
21	90	12.6	8491	15	US-10-133-013-260
22	90	12.6	8493	13	US-10-071-766-51
23	90	12.6	8503	15	US-10-037-270-130
24	90	12.6	8503	15	US-10-117-722-130
25	89.5	12.5	1071	10	US-09-933-767-118
26	89.5	12.5	1071	14	US-10-004-860-118
27	89.5	12.5	1071	14	US-10-023-282-118
28	89.5	12.5	1107	16	US-10-424-599-112896
29	89.5	12.5	9415	15	US-10-311-455-268
30	89	12.4	864	15	US-10-320-800-49
31	89	12.4	3522	18	US-10-425-115-19622
32	89	12.4	5262	18	US-10-425-115-19617
33	88.5	12.4	3370	17	US-10-437-963-83303
34	88.5	12.4	3966	17	US-10-437-963-83303
35	88	12.3	858	17	US-10-425-115-147531
36	88	12.3	1647	18	US-10-085-117-73
37	88	12.3	9934	15	US-10-085-117-73
38	87	12.2	3720	13	US-10-007-805-576
39	87	12.2	3720	15	US-10-076-622-576
40	87	12.2	3720	15	US-10-124-805-576
41	87	12.2	3873	17	US-10-437-963-41655
42	87	12.2	81440	17	US-10-451-467A-659
43	87	12.2	123192	14	US-10-175-523-71
44	86.5	12.1	776	16	US-10-424-599-23224
45	86.5	12.1	1181	16	US-10-415-134-3

ALIGNMENTS

RESULT 1
US-10-005-907-1
; Sequence 1, Application US/10005907
; Publication No. US20030166881A1
; GENERAL INFORMATION:
; APPLICANT: Union Chimique Belge, S.A.
; APPLICANT: Pirozzi, Gregory
; APPLICANT: Einstein, Richard
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAST
; TITLE OF INVENTION: ACTIVATION
; FILE REFERENCE: 053529-5005
; CURRENT APPLICATION NUMBER: US/10/005.907
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3762
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: